Porelike Morphologies in Amyloidogenic Proteins

A Thesis

Submitted to the Faculty

of

Drexel University

by

Matthew J. Voelker

in partial fulfillment of the

requirements for the degree

of

Doctor of Philosophy

September 2017

Acknowledgments

I would first like to thank my advisor, Dr. Brigita Urbanc, for all of her guidance throughout my journey as a graduate student. She has taught me valuable lessons that I will carry in my personal and professional life forever. I am indebted to Dr. Derya Meral for all of her help and wisdom throughout my graduate career and for being a wonderful friend. I also would like to thank Dr. Bogdan Barz for his teachings and his patience. In addition, I would like to thank the brilliant people who have been a part of my thesis advisory committee, for their insight and for providing motivation; Drs. Luis Cruz Cruz, Daniel Marenda, Jian-Min Yuan, Eric Brewe, Naoko Kurahashi Neilson, and Robert Gilmore.

Thank you to all of my friends who have shared this journey with me and helped ease the pain of being a graduate student for six long years, Ian Eaves, Matthew Mawhinney, Justin Bird, John Bridstrup, Kelly Douglass, Cindy Lin, Narayan Wong, and Vishal Kasliwal. Thank you Dylan Beery, Chris Khosrovani, and Mike Scorza for the adventures and friendship that often provided a much needed escape from my life as a graduate student, and to Meghan Jacobs and Scott Beseth for the life-long friendship.

I cannot possibly thank my parents enough for all of their love and motivation throughout my life. John Voelker and Loralee Hamel, you always encouraged me to be who I wanted to be, and I would not be here without you. Thank you to my big sisters Andrea and Suzanne Voelker, for all of your advice, for always listening, and for being a safety net that I can always rely on.

Finally, I would like to acknowledge my partner Sarah Gannon. Your love and unending support carried me forward when I was unable to carry myself. You are a constant source of inspiration.

Table of Contents

List of Tables

List of Figures

Abstract

Porelike Morphologies in Amyloidogenic Proteins
Matthew J. Voelker
Dr. Brigita Urbanc

Intrinsically disordered proteins (IDPs) have been linked to a variety of human diseases. The roles that IDPs play in physiological functions and disease pathology are frequently an enigma. Their disordered nature and structural complexity presents significant experimental and computational challenges, and makes IDPs difficult to study and characterize effectively. Soluble, low molecular weight (LMW) oligomers of the IDPs amyloid β-protein (Aβ) and α-synuclein (αS) have been hypothesized to be the primary neurotoxic agents in Alzheimer's Disease (AD) and Parkinson's Disease (PD) respectively, however their structure remains elusive. In this thesis, we take a varied computational approach in studies of Aβ and αS oligomers in order to probe and elucidate their structure.

Aβ oligomers have been observed to impair cognition in live rats and to negatively affect memory by hindering long-term potentiation in the hippocampus. Of the two predominant Aβ alloforms, Aβ40 and Aβ42, the latter is more strongly associated with AD. Here, we structurally characterize Aβ40 and Aβ42 monomers through pentamers via a multi-scale computational approach, wherein conformations derived by discrete molecular dynamics combined with an implicit–solvent intermediate–resolution protein model and amino acid-specific interactions (DMD4B-HYDRA) were converted into all-atom conformations and subjected to explicit-solvent MD. Unlike the initial DMD4B-HYDRA conformations, fully atomistic Aβ40 and Aβ42 trimers, tetramers, and pentamers form water-permeable pores, whereby the tendency for pore formation sharply increases with oligomer order and is the highest for Aβ42 pentamers. Our findings reveal an extraordinary ability of Aβ oligomers to form pores in pure water prior to their insertion into a membrane.

PD is characterized in part by the cerebral accumulation of α-synuclein, a 140 amino acids–long protein produced naturally in the body, into abnormal intracellular protein deposits in the brain.

Soluble αS oligomers in particular have been shown to be toxic to neuronal cell cultures *in vitro*. We here characterize the structure of αS oligomers computationally using the DMD4B-HYDRA approach. We vary the implicit-solvent parameter corresponding to the strength of electrostatic interactions (E_{CH}) to fine tune the solvent conditions in order to obtain an αS oligomer distribution consistent with experimental data. The population of αS oligomers is characterized by a monotonically decreasing propensity of monomers through septamers followed by an increase in octamers. We observe that αS forms water-permeable pores in all assembly states, with a propensity for pore formation that increases with oligomer order.

Previous studies have reported that both Aβ and αS oligomers form ion channels when embedded into a cellular membrane, which causes an abnormal ion flux and eventually leads to cell death. These observations form a leading theory for the mechanism behind the cytotoxicity associated with AD and PD, known as the ion channel hypothesis. The observations of porelike morphologies in our studies of LMW oligomers of Aβ and αS, which form in the absence of a membrane and increase in propensity with oligomer order, provides important insights and further support to the ion channel hypothesis of AD and PD.

Chapter 1: Introduction

Proteins are a fundamental aspect of all lifeforms and play a vital role in our everyday physiological processes, health and well-being. The focus of this thesis is on a specific class of proteins known as intrinsically disordered proteins (IDPs), that are by their nature disordered and difficult to characterize. Aberrant aggregation of IDPs has been linked to multiple human diseases, including Alzheimer's Disease (AD) and Parkinson's Disease (PD), which are associated with IDPs amyloid β-protein (Aβ) and α-synuclein (αS), respectively. Low molecular weight (LMW) oligomeric assemblies of these proteins inparticular have been posited as the primary cytotoxic agents in AD and PD. Characterizing LMW oligomers of Aβ and αS could thus provide crucial insights into the pathology of these diseases and may allow for the development of therapeutic treatments for those suffering from AD and PD. A primary hallmark of both AD and PD is the loss of neurons. One proposed method of cellular toxicity induced by Aβ and αS is the ion channel hypothesis, which posits that annular pores formed by assemblies of Aβ and αS can act as ion channels within the cellular membrane, resulting in an influx of calcium ions into the cell which can result in cell death. In this chapter, the fundamentals of protein structure, dynamics, and assembly are discussed, along with background information on the relevance of Aβ and αS in the context of the AD and PD pathology.

Experimental techniques lack the resolution to adequately characterize the structure and formation of LMW oligomers, which motivates the need for computational approaches. However, there are several challenges associated with modeling and simulating IDPs such as Aβ and αS. It takes considerable time for unfolded and spatially separated peptides to assemble into oligomers, and even more time for the resulting populations to reach a meta-stable state. Populations of Aβ and αS assemblies are heterogeneous, and it is thus important to ensure that the phase space available to these populations is adequately sampled. Explicit-solvent, fully-atomistic molecular dynamics (MD) provides for atomic level resolution when simulating protein folding and dynamics. However, MD requires numerical integration to simulate particle trajectories, which necessitates the use of signifi-

cant computational resources. MD thus cannot reasonably attain the time scales needed to observe oligomerization from unstructured and separated peptides. Discrete molecular dynamics (DMD), a more efficient version of MD, combined with an implicit-solvent and intermediate-resolution peptide model[1] (DMD4B-HYDRA) is able to overcome the computational obstacles associated with simulating assembly of unfolded and spatially separated monomeric peptides into oligomers. This approach allows for simulation times which are long enough for oligomer assembly to occur, and for the resulting populations to become meta-stable. The intermediate-resolution peptide model entails the use of four 'beads' to represent each amino acid. This representation relatively accurately models the protein backbone and accounts for the unique side-chain associated with each amino acid. Each amino acid retains its chiral character, and is allowed to form two hydrogen bonds (except proline). In order to take advantage of the strengths of MD and DMD, and to compensate for their respective weaknesses, these two approaches can be combined into a multi-scale approach. In this unique and novel computational method, DMD4B-HYDRA is used to create oligomer populations from unfolded and spatially separated peptides, which are then converted into fully-atomistic monomer and oligomer representations and subjected to explicit-solvent MD. Both MD and DMD play important and distinct roles in the field of protein physics, and these techniques, along with the multi-scale approach which combines the two computational techniques, are discussed in detail In Chapter 2.

The studies in this thesis applied the above computational techniques to simulations of Aβ and αS in order to gain insight into the nature and characteristics of folding and oligomer assembly, with the ultimate goal of providing valuable knowledge that contributes to the development of therapies for AD and PD. In Chapter 3, the multi-scale computational approach described above was applied in order to generate and structurally characterize a population of Aβ monomers through pentamers. In Chapter 4, a DMD4B-HYDRA study of αS folding and oligomer formation is presented. The most important result in both of these studies is the observation of porelike annular assemblies formed by Aβ and αS assemblies, the propensity for which increased with assembly size. To the best of our knowledge, these are the first studies to observe the formation of porelike annular oligomers of both Aβ and αS starting from spatially separated, unstructured peptides.

1.1 Protein Structure

Proteins are essential to all known forms of life. They play a vital role in a variety of biological functions, from immune system function to cellular communication. Although they are ubiquitous in life on earth, proteins in many ways remain an enigma. Proteins are encoded by DNA, and are thus subject to inherent variation. Protein levels can vary significantly between individuals, subgroups, and large populations as a whole[2]. Incorrect encoding of proteins occurs frequently and is largely benign, but this can result in abnormal and potentially even pathogenic behavior. Even correctly encoded proteins can be incredibly sensitive to environmental changes, which can disrupt their normal function and result in cytotoxicity. Several human diseases have been associated with the abnormal behavior of proteins, including Alzheimer's Disease (AD) and Parkinson's Disease (PD), among several others. Gaining insight into how proteins function, as well as how and why they can become pathogenic, is essential to our understanding of life itself.

The encoding of proteins by DNA can be thought of as a set of instructions for creating a chain of amino acids linked together by covalent bonds. 20 different amino acids occur in nature, and all proteins can be described by the sequence of amino acids that make up each protein, which is known as the protein's primary structure. Proteins can have a primary structure up to tens of thousand of amino acids long, and with 20 different amino acids, the number of possible variations in primary structure that can occur is effectively infinite.

1.1.1 Amino Acids

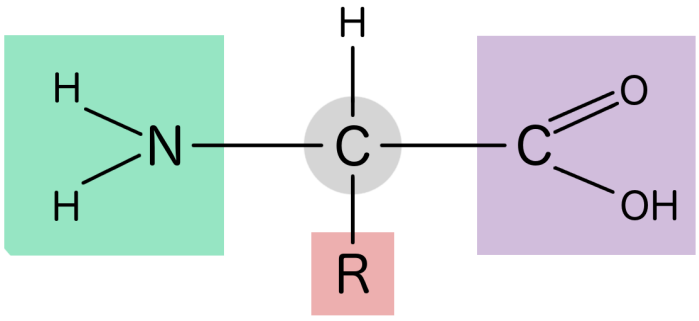

Figure 1.1: Amino acid structure.

Every amino acid is composed of an α-carbon (C_α), a carboxyl group (COOH, labeled C'), and an amino group (NH_2, labeled N), as displayed in Figure 1.1. At physiological pH, its is likely that the carboxyl group will be deprotonated (COO^-) and the amino group protonated (NH_3^+). Differentiation between amino acids occurs due to the presence of a unique side-chain ('R' in Figure 1.1) which is bonded to the central C_α atom. The side-chain determines if an amino acid is hydrophobic (non-polar) or hydrophilic (polar),and can carry a net charge. The side-chain thus determines how amino acids interact with other amino acids, water, and other solvents. The Kyte-Doolittle (KD) hydropathy scale[3], which is derived from physical and chemical properties of amino acid side-chains, represents one of the standard quantizations of the relative hydrophobicity of all 20 amino acids, shown in Table 1.1.

Table 1.1: Kyte-Dootlite Hydropathy values and charge at neutral pH.

AMINO ACIDS				
Symbol	Full Name	Abrv.	K+D	Charge @ pH = 7
I	Isoleucine	Ile	4.5	0
V	Valine	Val	4.2	0
L	Leucine	Leu	3.8	0
F	Phenylalanine	Phe	2.8	0
C	Cysteine	Cys	2.5	0
M	Methionine	Met	1.9	0
A	Alanine	Ala	1.8	0
G	Glycine	Gly	-0.4	0
T	Threonine	Thr	-0.7	0
S	Serine	Ser	-0.8	0
W	Tryptophan	Trp	-0.9	0
Y	Tyrosine	Tyr	-1.3	0
P	Proline	Pro	-1.6	0
H	Histidine[a]	His	-3.2	0/+1
Q	Glutamine	Gln	-3.5	0
N	Asparagine	Asn	-3.5	0
D	Aspartic Acid	Asp	-3.5	-1
E	Glutamic Acid	Glu	-3.5	-1
K	Lysine	Lys	-3.9	+1
R	Arginine	Arg	-4.5	+1

[a] Histidine is predominantly uncharged at neutral pH.

1.1.2 Peptide Bonds and Backbone Geometry

Amino acids are linked together into a peptide chain by covalent bonds known as peptide bonds. Peptide bonds form through a process known as condensation, visualized in Figure 1.2. A reaction occurs between the carboxyl group of one amino acid and the amino group of the neighboring amino acid, resulting in a covalent bond between the carboxyl carbon and amino nitrogen. As a byproduct of this reaction, the OH from the carboxyl group and one of the amino group hydrogens separate from their respective amino acids and come together to form a water molecule. This reaction occurs between pairs of amino acids to form a polypeptide chain, and one or more polypeptide chains form a protein. Amino acids that are part of a protein can be referred to as residues, and the continuous connection of peptide bonds which forms a complete protein chain is known as the protein's backbone. Proteins have an explicit directionality due to the asymmetric nature of peptide bond formation, as there will always be an amino group and a carboxyl group at opposite ends of a protein chain which do not form peptide bonds. The typically protonated amino group at the beginning of a protein chain is referred to as the N-terminus, and the typically deprotonated carboxyl group at the end of the chain is referred to as the C-terminus.

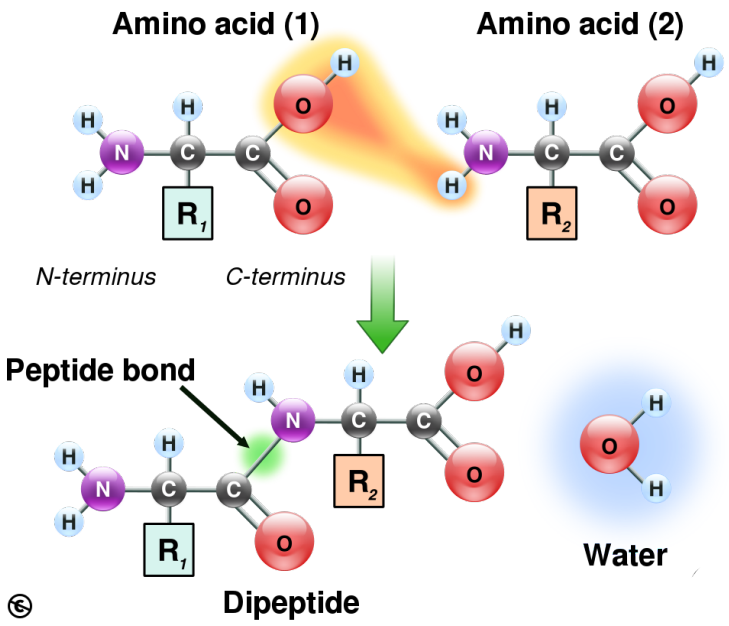

Figure 1.2: Peptide bond formation.

Peptide bonds are fairly rigid structures, which imposes inherent constraints on the flexibility of the protein backbone, and therefore the protein as a whole. The orientation of the backbone can be described with three geometric planes, whose orientation can be defined by the dihedral angles, ϕ, ψ, and ω (visualized in Figure 1.3). The amino acid structure and the rigidity of peptide bonds inhibit the rotational freedom of backbone atoms, and therefore the dihedral angles for a given set of neighboring backbone atoms do not occupy every region of the conformational phase space defined by the dihedral angles ϕ and ψ (referred to as Ramachandran plots) with equal propensity. The dihedral angle ω specifically is limited to either 180° or 0°, with the former being the significantly more typical orientation.

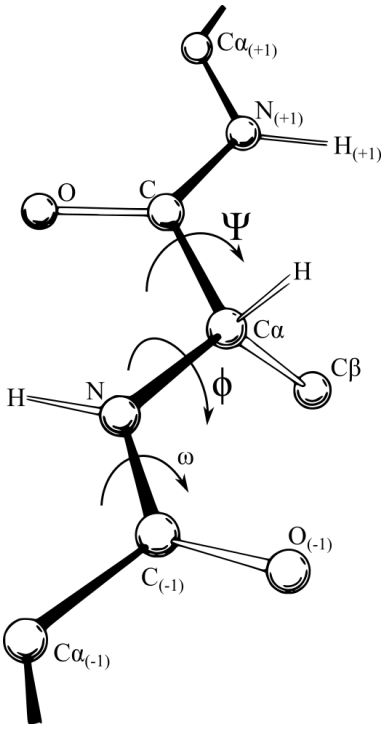

Figure 1.3: Backbone dihedral angles. Image by Dcrjsr, vectorised by Adam Redzikowski (Own work) [CC BY 3.0 (http://creativecommons.org/licenses/by/3.0)], via Wikimedia Commons.

1.1.3 Secondary Structure

Ordered, folded structures along local regions of the primary structure are referred to as secondary structure. The three most common types of secondary structure are β-strands, turns, and α-helices,

which can be thought of as different orientations of the protein backbone and often involve intra-peptide hydrogen bond formation. β-strands are characterized by several consecutive residues in an extended conformation with an explicit orientation. β-strands often form in laterally hydrogen bonded pairs in a parallel (same orientation) or anti-parallel (opposite orientation) manner, such as in a β-hairpin conformation (Figure 1.4A). Larger groups of connected β-strands can form a β-sheet, again in a parallel or anti-parallel orientation. α-helices are spiral conformations formed by a string of amino acids, within which hydrogen bonds between amino and carboxyl groups (separated by three to fours residues in the sequence) stabilize the helical structure (Figure 1.4B). Turn structures are simply a change in direction of the protein chain on a local scale, and are often in a 'U-turn' shape. A fourth descriptor of the secondary structure, random coil, exists to describe a lack of the above mentioned well-defined secondary structures. The different types of secondary structure arise from interactions and bond formation between backbone atoms and do not explicitly involve residue side-chains, although the presence of side-chains can affect the secondary structure propensity of individual amino acids in short peptides[4–7]. The side-chain of glycine, for example, is simply a hydrogen atom, which makes glycine the most flexible amino acid in the context of higher order protein structure formation. Proline's unique side-chain forms a ring structure by bonding the C_β atom to its own amino group, giving proline a notably rigid structure and resulting in an inability to form two backbone hydrogen bonds. Proline is known as a structure breaker, and consequently it is often found in turn structures and at the beginning of α-helices, rather than in the middle.

1.1.4 Tertiary and Quaternary Structure

The highest order of protein structure that can be ascribed to a monomeric peptide is the three-dimensional atomic coordinates of all the residues in a peptide, known as its tertiary structure. The tertiary structure is composed of all the local secondary structure along the primary sequence as well as any random coil-like regions which are not in any secondary structure orientation (Figure 1.5A). Tertiary structure is stabilized by local secondary structures, backbone hydrogen bonds, and side-chain-side-chain interactions. When two or more peptides assemble into an oligomer, through inter-molecular hydrogen bonds or other interactions, the arrangement of peptides in the resulting oligomer

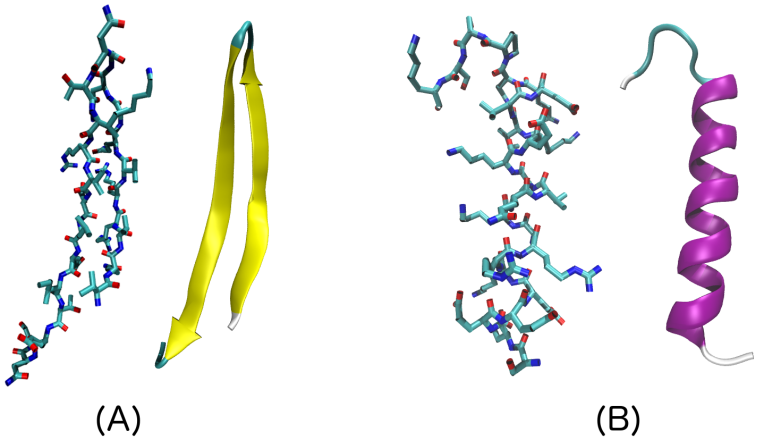

Figure 1.4: Common Secondary Structures. (A) β-hairpin conformation in two different representations; (B) α-helix conformation in two different representations. PDB (ID: 4R80) provided by the Protein Data Bank[8].

is described by its quaternary structure (Figure 1.5B). Quaternary structure can be thought of as multiple tertiary structures that have become interconnected. In globular (sphere-like) proteins, the tertiary and quaternary structure is often characterized by a core composed of the hydrophobic residues in the sequence with hydrophilic residues residing on the surface around this hydrophobic core. As will be discussed later in this Chapter, a functional monomeric protein can adopt a distinct tertiary structure that is associated with its physiological function, referred to as the protein's native state. The native state of some proteins is an oligomeric form, and thus can be described by both the tertiary structure of the individual peptides, and the quaternary structure of the oligomer.

1.2 Protein Folding: The Native State

The native, biologically functional, folded state of a protein is only one of an astronomically large number of possible conformations that a protein can take on. Even relatively short proteins can adopt complex structures that are sensitive to a number of environmental conditions. A protein of sufficient length could spend more time exploring all of the conformations available to it, given sufficient energy, than the current age of the universe[9]. Yet there are countless proteins in the human body that manage to reliably fold into their native state, sometimes in a matter of micro-

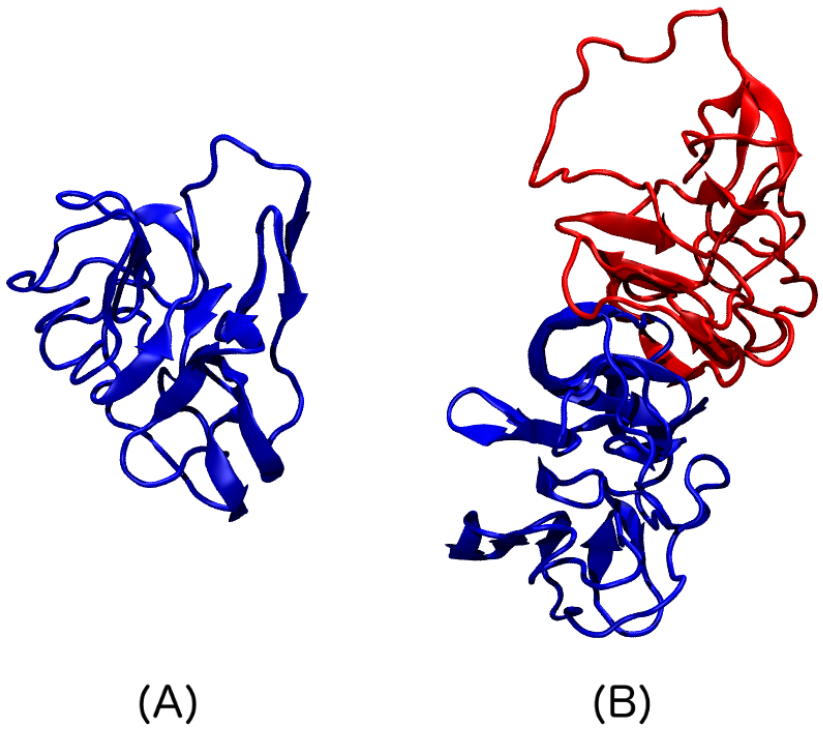

Figure 1.5: Tertiary and Quaternary Structure. (A) Tertiary structure of an α-synuclein monomer with multiple β-strands (arrows); (B) Quaternary structure of an α-synuclein dimer, with interconnected β-strands from each peptide.

to milliseconds. These seemingly contradictory observations inherently lead to a scientific question which remains relevant today; how do proteins fold into their native state so quickly and relatively reliably when it is only one possible conformation among so many? Francis Crick addressed this 'central dogma' in a 1957 lecture, stating that all of the information needed to discern how a protein will fold is contained in its primary structure[10;11]. Although the actual physical process of encoding a protein is certainly more complex, it can be thought of simply as a set of instructions for which amino acids should go in which order. This is all of the 'information' that is present in a given fragment of DNA responsible for protein encoding, and therefore this is the only information that is passed on to the protein. Francis Crick thus argued that the information contained in the primary structure must therefore be sufficient for a protein to bypass the countless conformational possibilities available to it and fold into its native state. Christian Anfinsen expanded on this idea in 1973 by proposing three conditions which need to be met in order for a protein's native structure to be determined solely by its primary structure[12]. Importantly, Anfinsen also introduces conditions on the protein's environment. Dubbed 'Anfinsen's thermodynamic hypothesis,' it states that: the native state must be unique in that no other conformation has a comparable free energy, the native state must be stable enough to avoid being significantly perturbed by changes in the protein's environment, and the protein's journey through the free energy landscape must not require complex conformational changes to take place in order to reach the native state[12].

1.2.1 Free Energy Landscapes

The predominant approach which arose to answer the scientific question of how a protein 'knows' how to attain its native state, revolves around the thermodynamic concept of free energy landscapes, which was formulated for proteins by Bryngelson and Wolynes[13]. By selecting certain physical parameters which can be measured or calculated for all the possible conformational states a protein can adopt, one can create a multi-dimensional conformational phase space. Each combination of possible values for the chosen parameters, or reaction coordinates, can be described by its thermodynamic free energy, most commonly its Gibbs free energy. Each point in the conformational phase space is given a depth which corresponds to its free energy, or its change in free energy relative to some initial

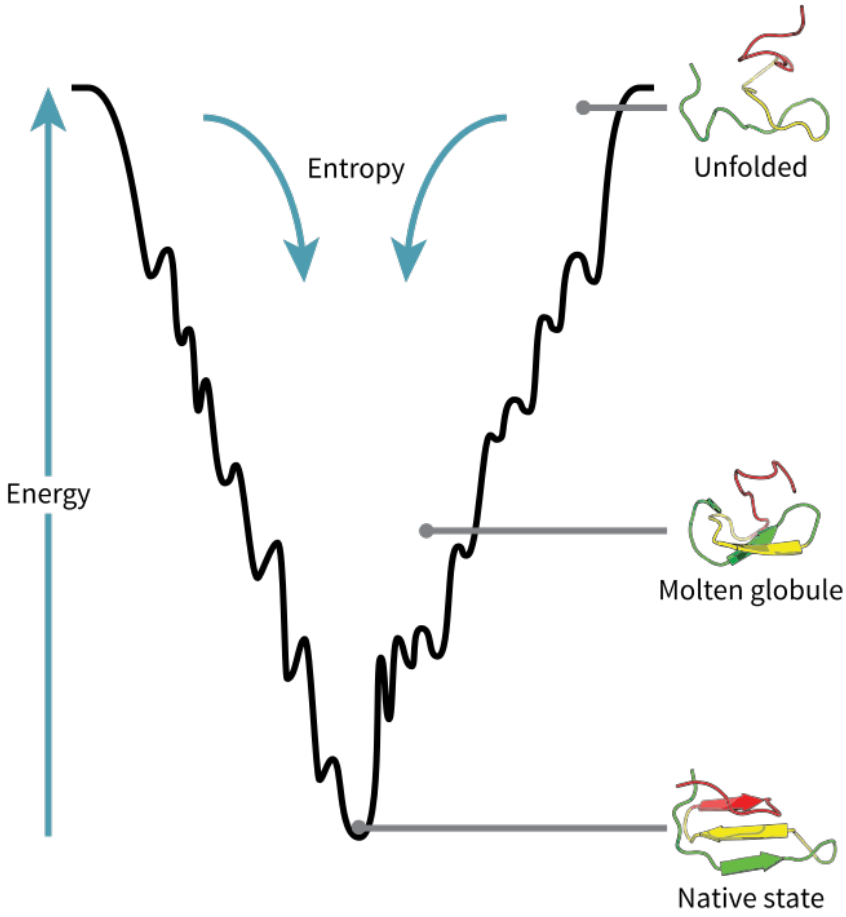

Figure 1.6: Illustration of a free energy landscape. The peptide starts at the top of the landscape in an unfolded conformation. The pathway down the landscape corresponds to a lowering of its free energy as it folds into its native state. Image by Thomas Splettstoesser (www.scistyle.com) (Own work) [CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons.

conformation, creating a landscape. Nature inherently favors states with the lowest free energy, as these are the most energetically stable[14]. Bryngelson and Wolynes ascribed this idea to proteins with the 'principle of minimal frustration,' which states that the native structure of an amino acid sequence is inherently very stable, otherwise that sequence would not have been naturally selected to perform some biological function[13]. By this logic, a protein would be expected to traverse its free energy landscape by folding in a manner that lowers its free energy, eventually folding into its native state by reaching a reliably attainable and stable free energy minima. A visual representation of this process is displayed in Figure 1.6.

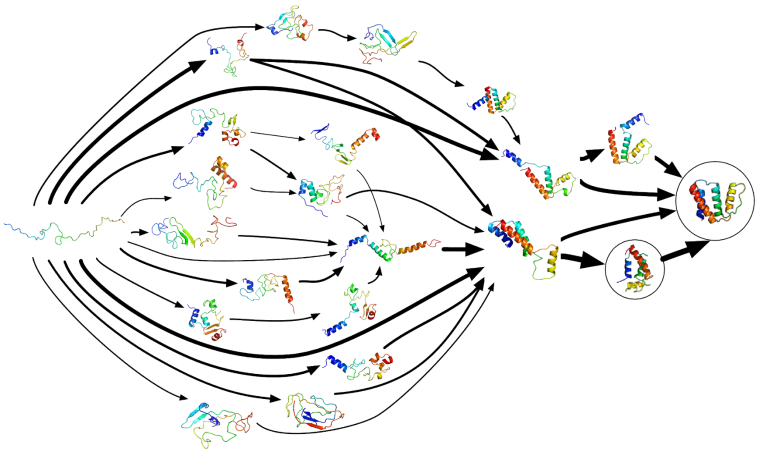

Figure 1.7: Illustration of multiple folding pathways from unfolded to native state. Image by Vincent Voelz [CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons.

In practice, the pathway from unfolded to native state for many proteins is not as simple as what has been historically theorized. Free energy landscapes are often very complex and can feature multiple folding pathways[15], an example of which is displayed in Figure 1.7. Experimental studies have shown that some proteins will fold, unfold, and refold multiple times before reaching their native state[16;17]. Recreating the precise *in vivo* environment in which proteins fold presents significant and in some aspects currently insurmountable challenges, both experimentally and computationally. Furthermore, an entire class of proteins known as intrinsically disordered proteins lack a well-defined native fold altogether, and are characterized by significant structural disorder in large sections of the primary structure. Studying this class of proteins requires special consideration in order to account

for the inherent unpredictability of their structural dynamics.

1.2.2 Driving Forces Behind Protein Folding

The physical forces and interactions which drive the folding and dynamics of proteins have to be modeled and accounted for when studying the behavior of proteins. The various environments that proteins in nature inhabit can have a significant influence on protein folding. For example, cellular environments can be very crowded, and an entire field of research is devoted to studying the effects of volume exclusion and the influence of macromolecular crowders and osmolytes on protein folding[18;19]. However, the most ubiquitous environmental solvent that proteins exist within is water. Modeling water is deceptively complex, and extensive effort has been put forth over the last several decades to accurately model the effects of water on protein folding.

The hydrophobic nature of certain amino acids, and the presence of a significant amount of hydrophobic (non-polar) residues in many proteins, results in hydrophobicity being a dominating force of protein folding. Table 1.1 displays the hydrophobicity of all 20 naturally occurring amino acids as formulated by the KD hydropathy scale. Hydrophobicity and hydrophilicity can be modeled as a physical force of repulsion or attraction, respectively, between residues and between residues and water. Consequently, hydrophobic residues tend to cluster together in order to shield themselves from water. Conversely, hydrophilic residues tend to exist on the surface of a peptide's tertiary structure or sometimes can be found fully exposed to the solvent, such as in the flexible and solvent exposed N-terminus of Aβ42 oligomers[20;21].

Electrostatic attraction and repulsion between side-chains influences protein structure through coulomb forces and the formation of salt bridges. The strength of electrostatic forces is highly influenced by the solvent, as high salt concentrations can screen electrostatic interactions and reduce the effective interaction range of coulomb forces between side-chains. Salt bridges formation is initiated by attraction between amino acid side-chains which are oppositely charged. A salt bridge is then created when a hydrogen bond forms between the carboxyl oxygen (which is not already double-bonded) in the negatively charged residue and the amino nitrogen in the positively charged residue. Salt bridges can help to stabilize tertiary and quaternary structure, even if the conformation

is not entropically favorable, such that in most cases salt bridges are exposed to the solvent.

In addition to hydrophobicity and explicit electrostatic interactions, van der Waals (VDW) forces contribute to protein-protein and protein-solvent interactions as well[22]. Although VDW forces are weaker, they are still relevant at small inter-atomic distances. VDW forces account for interactions that arise from induced dipoles, resulting in both attractive and repulsive forces. These forces are modeled classically using the Lennard-Jones (LJ) potential[23]. The functional form of the LJ potential [eq. 1.1] is displayed and plotted in Figure 1.8 for a given inter-particle distance 'r'.

$$u(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{1.1}$$

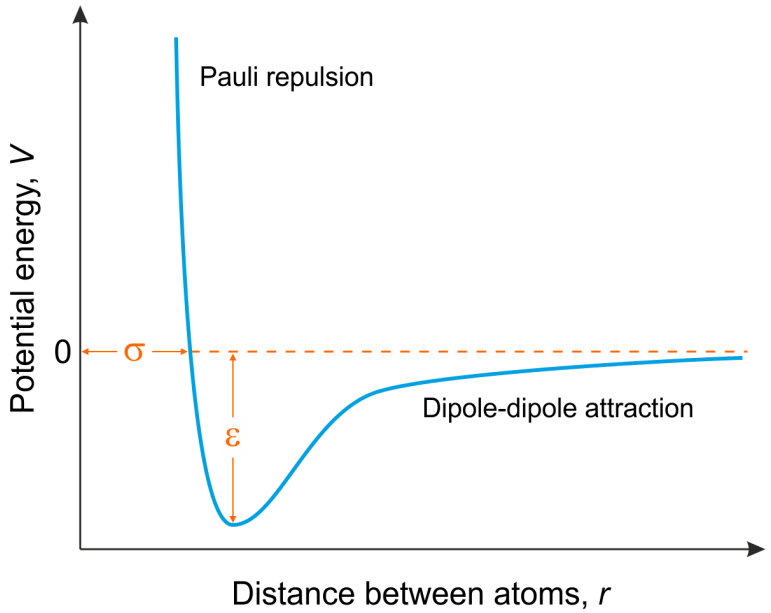

Figure 1.8: The Lennard-Jones Potential. Image by Generalic, Eni. "Lennard-Jones potential." Croatian-English Chemistry Dictionary & Glossary. 29 Aug. 2017. KTF-Split. Accessed: 8 Sep. 2017. (glossary.periodni.com)

The deepest point in the potential is -ϵ, as defined in eq. 1.1. Repulsive forces dominate to the left of this deepest point, whereas attractive forces dominate on the right. For values of $r < \sigma$, the potential quickly approaches $+\infty$ in order to prevent two particles from occupying the same point in space simultaneously. For $r >> \sigma$, the potential dissipates towards zero, as the range of VDW forces is very small[23]. In course-grained approaches to protein simulations, such as the DMD method which will be discussed in Chapter 2, the VDW potential is replaced by simple hard-sphere

repulsion, wherein the potential is set to $+\infty$ for inter-particle distances which are smaller than the sum of the interacting particles' radii, and zero elsewhere.

1.3 Intrinsically Disordered Proteins and Human Diseases

Intrinsically disordered proteins (IDPs) are a distinct class of peptides characterized by a lack of ordered secondary and tertiary structure along a part or all of their amino acid sequence. IDPs seem to oppose the notion that the native conformation of a protein is structurally well defined and unique. If the native state of a particular protein is one in which a significant portion of the primary structure is in a random coil configuration, than it is difficult to reconcile how any particular conformation could occupy a unique, reliably attainable location in the free energy landscape, as suggested by Anfinsen's thermodynamic hypothesis[12]. And yet it has been established that significant disorder can be found *in vitro* in proteins with known biological significance[24–26]. Furthermore, a study on the IDP ΔK280 Tau found that the it adopts a relatively small number of conformations whose energy does not vary drastically[25;27]. It is not unreasonable, therefore, to propose that proteins whose observed conformations are intrinsically disordered can also serve a biologically functional role in the body[25;26;28]. Studying IDPs in the context of their role as functional proteins is an important component in our understanding of how proteins drive the biological processes that are essential to life. However, there is another side to the nature of IDPs which is also of great importance and relevance to humanity; their significant connection to the pathology of several human diseases.

1.3.1 Amyloid β-Protein and Alzheimer's Disease

Alzheimer's Disease (AD) is a neurodegenerative disorder which is the leading cause of dementia in the elderly[29]. Autopsies of AD brains revealed the presence of intra-neuronal neurofibrillary tangles and extracellular amyloid plaques among neurons which had become unviable prior to death[29;30]. A primary component of these amyloid plaques are fibrillar aggregates of two proteins which are common in the brain, the 40 and 42 amino acids-long alloforms of amyloid β-protein (Aβ)[29;31;32], which established their relationship to the pathology of AD. Adding evidence to the connection between these Aβ peptides and AD was the discovery of inherited genetic mutations, some of which

increase Aβ production while others affect the ratio of Aβ42 to Aβ40 in the brain, with both scenarios resulting in aggressive forms of AD that form on average much earlier in life than is observed in the typical AD pathology[29;33;34].

Aβ40 and Aβ42, along with several other alloforms of varying lengths, are all cleaved from amyloid precursor protein (APP) in the brain by two protein complexes, β-secretase and γ-secretase[35;36]. Aβ40 and Aβ42 are identical to each other with the exception of the C-terminus of the primary structure, at which Aβ42 has an additional two hydrophobic residues (I41-A42). In spite of this nominal difference, Aβ42 is significantly more prone to aggregate than its shorter counterpart[37;38]. Genetic mutations that shift the ratio of Aβ42:Aβ40 towards the former and result in aggressive AD pathologies have implicated Aβ42 as the more neurotoxic alloform. Due to the discovery of Aβ fibrils in AD brains, these large aggregates were believed to be a direct cause of neurotoxicity in the AD pathology. This view has shifted, however, as evidence began to accumulate implicating soluble, LMW oligomers of Aβ as the primary cytotoxic agent in AD[39–45].

Aβ40 and Aβ42 MD[46–67] and DMD[21;51;68–74] studies of Aβ40 and Aβ42 have probed and elucidated the nature and behavior of these two peptides. Urbanc *et. al.* were able to tune the parameters of the DMD4B-HYDRA force field in order to match experimentally derived oligomer distributions of WT Aβ40 and Aβ42, as well as their arctic mutants. Meral *et. al.* used DMD4B-HYDRA to explore the folding and aggregation of N-terminally truncated Aβ40 and Aβ42, which have been shown experimentally to seed Aβ aggregation and contribute to the toxicity associated with these peptides[75;76]. Their results suggested that the A$\beta 40_{3-40}$ and A$\beta 42_{3-42}$ alloforms had more flexible and solvent-exposed N-termini than full-length Aβ40 and Aβ42, which may explain their increased propensity to oligomerize and the toxicity associated with these shortened Aβ variants[73]. Barz *et. al.* performed simulations which combined DMD4B-HYDRA and fully-atomistic explicit-solvent MD, using the multi-scale computational approach described in Chapter 2, on Aβ40 and Aβ42 monomers and dimers in a precursor to the study in Chapter 3 of this thesis[60]. They observed that, similar to DMD4B-HYDRA simulations, the N-terminal region of Aβ42 dimers was more flexible and solvent-exposed than Aβ40 dimers, and that dimers of these two species were more

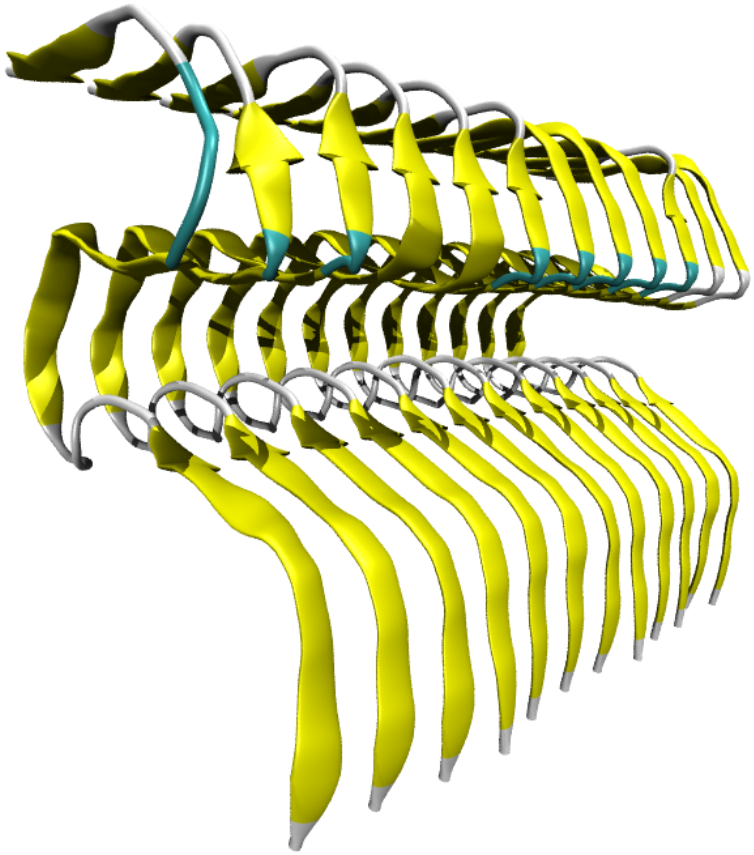

Figure 1.9: Model of Aβ42 fibrillar structure. PDB (ID: 2MXU) provided by Xiao *et. al.* and the Protein Data Bank[8;64]

distinct from each other than monomers of Aβ40 and Aβ42[60]. Xiao *et. al.* used solid-state NMR data to construct a fully-atomistic model for the Aβ42 fibrillar structure, which displayed a triple parallel β-sheet structure that can be seen in Figure 1.9[64]. The relevance of LMW oligomers to the pathology of AD motivates the need for their study and characterization. In Chapter 3 of this thesis, the same multi-scale approach used by Barz *et. al.*[60] was utilized in order to simulate Aβ40 and Aβ42 monomers and LMW oligomers up to and including pentamers.

1.3.2 α-Synuclein and Parkinson's Disease

Parkinson's Disease (PD) is a chronic degenerative disorder which affects motor neurons in the brain, resulting in severe impairments of the sufferer's ability to control their limbs, as well as the characteristic muscle spasms. Unlike AD, PD does not affect memory or directly impair cognition, although it

affects a patient's ability to speak. The PD pathology is characterized by intracellular protein aggregates known as Lewy bodies, which are also associated with a separate neurodegenerative disorder known as Lewy body dementia. The principle component of Lewy bodies is α-synuclein (αS), a 140 amino-acids long IDP. Similar to the relationship between Aβ and AD pathology, familial genetic mutations and enhanced expression of αS has linked this peptide to the pathogenesis of PD[77–82]. Oligomeric species of αS, rather than the fibrillar aggregates which are found in Lewy bodies, have been put forward as the primary neurotoxic agent behind the death of motor neurons in PD[83–88]. The non-Aβ component (NAC) of the αS primary structure has been shown to both seed and form the structural core of neuronal plaques associated with AD. As the NAC is derived from full-length αS, the name NACP (NAC-precursor) was given to αS prior to its current designation[89;90]. The NAC has since been shown to also be essential to the formation of the αS fibrils found in Lewy bodies and to the cellular toxicity associated with the PD pathology[91–94].

Unlike Aβ40 and Aβ42, full-length αS has not been subjected to DMD simulations prior to the study in Chapter 4, to the best of our knowledge. Dedmon *et. al.* used a combination of paramagnetic relaxation enhancement NMR spectroscopy to apply distance constraints to ensemble MD simulations, and their results suggest that interactions between the highly charged C-terminal region of αS and the NAC contributed significantly to the native peptide structure[95]. There is experimental evidence that suggests αS folds into an α-helical tetramer that resists aggregation[96], and enhanced sampling explicit-solvent MD simulations by Rossetti *et. al.* of monomeric N-terminal acetylated αS observed an amphipathic α-helical structure in the N-terminal region of the primary structure[97]. In Chapter 4 of this thesis, DMD4B-HYDRA was utilized to undertake, to our knowledge, the first ever computational study of αS assembly starting from spatially divided, unfolded monomers.

1.3.3 The Ion Channel Hypothesis

The exact mechanism of oligomer-mediated toxicity in AD and PD remains to be clearly elucidated. One of the leading hypotheses states that oligomers of Aβ and αS form annular pore structures which get embedded into cellular membranes and critically disrupt neuronal homeostasis by acting as ion channels for Ca^{2+} ions[83;98–113]. Indeed, calcium dysregulation has been posited to play a role

in both the cellular degeneration that is central to the PD pathology[114;115], and in the disruption of signaling pathways in AD, potentially resulting in the memory and learning impairments that characterize the disease[116]. Studies *in vitro* have shown that assemblies αS and Aβ can form ion channels when inserted into a lipid bilayer, with the latter having also been shown to induce an influx of calcium ions (Ca^{2+}) into the cell, resulting in the loss of cell viability[102;117;118]. However, the current resolution of experimental imaging techniques does not allow for the direct study of the fine structure of annular pores or the precise dynamics of their formation. Theoretical models of Aβ ion channels have been constructed and studied previously with MD[59;109;119–121], and these 'β-barrel' models were characterized by a significant amount of ordered structure and cylindrical symmetry. However, explicit-solvent MD simulations alone are not able to simulate long enough time-scales to observe the formation of porelike morphologies from initially unfolded and spatially separated peptides. Computational approaches which allow for both the observation of porelike assembly formation, and the subsequent pore structure, can provide important insights into the ion channel hypotheses of AD and PD.

Chapter 2: Computational Approaches to Simulating Proteins

In this Chapter, descriptions of classical molecular dynamics combined a fully-atomistic protein representation in explicit solvent, and an efficient variant of molecular dynamics called discrete molecular dynamics combined with a coarse-grained protein model and implicit solvent, are provided. These computational approaches are utilized in the studies in this thesis in order to simulate and characterize the IDPs Aβ40 and Aβ42 and α-synuclein. By their nature, IDPs occupy a complex conformational landscape. In order to effectively sample this landscape, our approach entails generating an ensemble of conformations in order to gather significant statistics for analysis[122]

2.1 Fully Atomistic Molecular Dynamics with Explicit Solvent

Molecular dynamics (MD) simulations can model the folding and dynamics of peptides with atomic level resolution. In MD, atomic interactions and collisions are governed by classical forces which are used to create continuous potentials (force field). This allows for integration of Newton's laws of motion in order to advance particle trajectories. Fully-atomistic (or all-atom) MD representations model every atom of each amino acid, including the unique side chains. Bonds between atoms and physical dimensions are modeled as geometric constraints and interparticle potentials rather than explicit structures. Explicit solvent models are often used in conjunction with MD in order to simulate an aqueous environment. Fully-atomistic MD allows for one of the most realistic representations of peptides and their behavior in a solvent that we have today. However, this resolution comes at a computational cost. Protein aggregation from unfolded and disassociated peptides may occur on time scales that would require unreasonable real-world simulation times, especially for larger proteins such as the α-synuclein or tau.

The MD study in this thesis (Chapter 3) utilized the GROMACS 4.9.5 package to generate and run simulation trajectories, which entails several steps[123–126]. First, a force field is chosen, which in our study was the OPLS-AA force field[127–130], whose parameters are used to build a model of

the protein using a conformation supplied by the Protein Data Bank (pdb)[8]. A simulation box is constructed around the peptide which can vary in shape and boundary behavior. The various ensembles in statistical mechanics can be applied to the simulation box depending on which physical properties one wishes to keep constant. The NPT ensemble was selected for our study, utilizing thermal coupling with velocity rescaling and a Berendsen barostat to maintain constant pressure and temperature[131]. The chosen water model is then used to generate explicit water molecules around the protein, along with enough Na^+ and Cl^- ions necessary to keep the simulation box electrically neutral overall. Once the simulation box is created and the protein system solvated, the system is equilibrated with the particle interactions turned off using a method of steepest descent until the potential energy of the system converges. In this step, particles are given random velocities pulled from a generated Maxwell distribution. Next, the peptide is constrained and the solvent is subjected to MD in order for water molecules it settle around the peptide in the most favorable configuration. Trajectories are then ready for production runs with all particle interactions implemented and the protein conformation unconstrained.

2.1.1 Force Field and Explicit Water Model

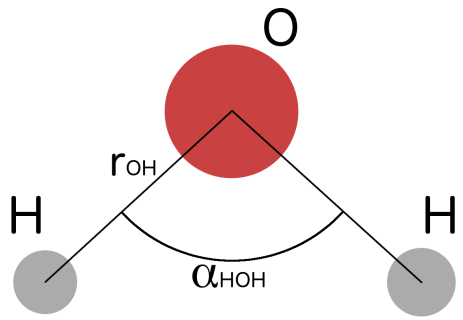

Figure 2.1: Three-point water model. In the TIP3P water model, r_{OH} = .9572 and α_{HOH} = 104.52°, and partial charges are applied to each atom (q(O) = -.0834, q(H) = .417)[127].

The choice of force field and water model can have a substantial effect on protein folding and dynamics, as evidenced by computational studies which compare results from simulations using different force fields and water models[7;60;132–134]. Force fields are collections of parameters and

equations which govern the evolution over time of bond lengths, angles, and torsions, as well as Lennard-Jones potentials and other interactions between atoms. These parameters can vary among MD studies depending on the force field. We have chosen for the all-atom MD study in Chapter 3 of this thesis the OPLS-AA (Optimized Potentials for Liquid Simulations-All Atom)[127–130] force field, which was parameterized based on quantum chemical calculations and experimental solvent data[130]. OPLS-AA was generated in conjunction with the TIP3P (Transferable Intermolecular Potential 3P) water model[127], a three-point model for water molecules, as seen in Figure 2.1. Each atom is assigned a partial charge in order to create a polar water molecule. TIP3P is one of the most frequently utilized water models in explicit solvent MD simulations. We pair the OPLS-AA force field with the TIP3P water model for the fully atomistic study of Aβ40 and Aβ42 monomers and oligomers in Chapter 3 of this thesis.

2.2 Discrete Molecular Dynamics

Discrete molecular dynamics (DMD) seeks to lessen the computational expenses associated with MD simulations, allowing for millisecond time scales in simulations of large systems with relative ease, whereas MD can only achieve such time scales by using immensely powerful and specialized supercomputers[135]. Rather than defining a continues potential to govern particle interactions and thus requiring numerical integration to advance particle trajectories, as in MD, the force-field in DMD is made up of pairwise, spherically symmetric square well potentials[136] with explicit short-range interaction distances (Table 2.1). The size of these potential wells are defined by an explicit interaction distance, and the depth corresponds to energies that can be positive or negative depending on the specific interaction. Periodic boundary conditions are applied to the edges of the cubic simulation box.

DMD is an event-driven form of MD; particle positions and velocities are only tracked and calculated when collisions occur. Both inelastic and collisions are modeled and the appropriate corresponding conservation laws (kinetic energy, momentum, and angular momentum) are applied. Between collisions, particles move at a constant velocity along a straight paths. The efficiency of the DMD algorithm scales as Nln(N), where N is the number of atoms in the simulation[136;137]. The

peak efficiency of the DMD algorithm is inversely proportional to the number of discontinuities in the potential. An implicit solvent with amino acid-specific interactions is instituted along with a four-bead amino acid model combined with DMD (DMD4B-HYDRA) in order to reduce the number of particles that are simulated, which adds to the efficiency of the DMD algorithm.

Table 2.1: Square well potentials in DMD between two atoms, A and B. The interatomic distance between atoms A and B is the variable r.r_i^{AB} represents a step in the potential with associated potential energy u_i^{AB}, r_{min}^{AB} is the hardcore collision distance, and r_{max}^{AB} is the maximal range of interaction, which is defined by the user[1].

Interatomic distance	Interaction pair potential
r_i^{AB} <r <r_{i+1}^{AB}	u_i^{AB}
r <r_{min}^{AB}	u = ∞
r <r_{max}^{AB}	u = 0

2.2.1 The DMD4B-HYDRA Approach

The DMD4B-HYDRA approach has been successfully applied in several computational studies of protein folding and aggregation. Studies of WT Aβ40 and Aβ42 proteins and their naturally occurring variants[21;51;68–74] were discussed in Chapter 1. DMD4B-HYDRA is not limited to the simulation of amyloidogenic proteins, however, as it has also been applied in studies of the non-amyloidogenic proteins stefin B[138] and pig gastric mucin[139]. The DMD4B-HYDRA approach was utilized in Chapter 4 of this thesis in a study on the folding and aggregation of αS, a 140 amino acids-long intrinsically disordered amyloidogenic protein, in order to be able to observe and characterize αS oligomer assembly, which would not be feasible with explicit-solvent MD.

Four-bead model

The efficiency of the DMD approach can be further enhanced by combining DMD with an intermediate resolution course-grained amino acid model, which depicts fewer atoms than an all-atom simulation while still retaining the necessary information to represent a sequence's behavior. In this model, the backbone is modeled quite precisely, amino acids retain their chiral character, and each amino acid, except proline, can form up to 2 backbone hydrogen bonds. In order to retain an accurate portrayal of the geometric properties of the peptide's backbone, the model requires at

least three atoms, called 'beads' in this model. These beads correspond to the atoms comprising the amino group, N, the α-carbon group, C_α, and the carboxyl group, C. We apply a four-bead model[1] by using a fourth bead, C_β, to represent the unique side-chain associated with each amino acid (Figure 2.2). Glycine amino acids are represented by only three beads, as its side chain is only a single hydrogen atom. The model incorporates permanent bonds between beads, wherein beads can assume any distance within the minimum and maximum bond length. Constraints between beads which never form permanent bonds in part define the backbone geometry.

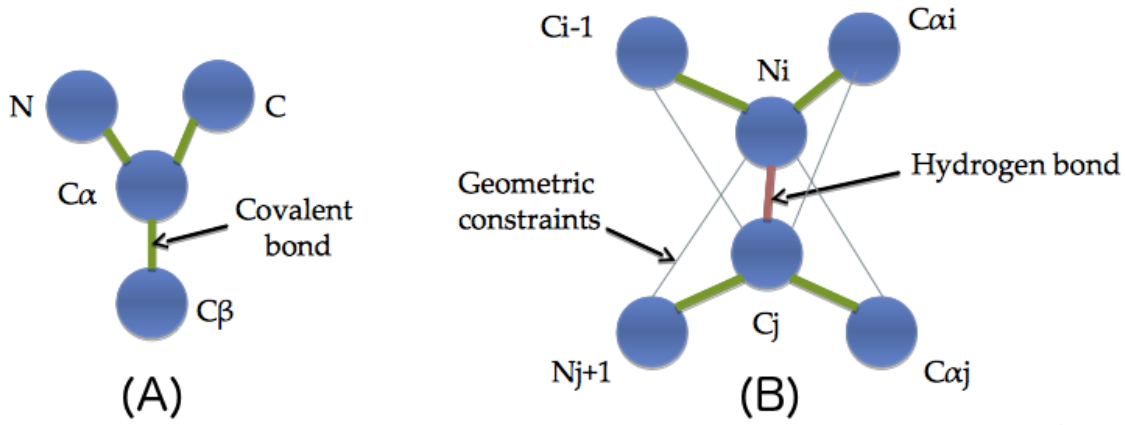

Figure 2.2: (A) Four-bead amino acid model; (B) Backbone hydrogen bonding in the four-bead model

The four-bead representation does not contain any oxygen atoms, and thus implements hydrogen bonding between the amide group of the i_{th} amino acid, N_i, and the carbonyl group of the j_{th} amino acid, C_j. The beads on either side of the N_i, and C_j atoms are subject to distance constraints via auxiliary bonds, which are necessary to accurately model the directionality of hydrogen bonds. The potential energy associated with these auxiliary bond distances is presented in Table 2.2[1]. A hydrogen bond will only form if all six beads (N_i, C_j, and their neighbors) are at distances and orientations which are energetically accessible to the system, i.e. if there is enough kinetic energy to overcome the inherent potential energy barrier associated with hydrogen bond formation. Thermal fluctuations can alter these distances, and if maintaining the hydrogen bond becomes energetically unfavorable, it will break. The parameters which define the distance constraints in the four-bead model and hydrogen bond geometry have been reported previously[1], and

Table 2.2: Potential energy associated with hydrogen bonding in the four-bead model. ϵ_{HB} is the potential energy associated with a hydrogen bond[1].

Potential Energy	Auxiliary Bond Distance
ϵ_{HB}	d_{min} <d <d_o
$\epsilon_{HB}/2$	d_o <d <d_1
0	d_1 <d <d_{max}
∞	else

are calculated from distributions of experimentally-derived measurements of $\sim$ 7700 folded proteins with known crystal structures[8], or chosen following the standard knowledge of the geometry of the peptide backbone[140]. Intra-molecular hydrogen bonds cannot form between amino acids that do not have at least three amino acids between them. We account for only inter- and intra-peptide backbone hydrogen-bond interactions, which are advantageously not amino acid-specific, and attempt to capture the backbone geometry with this simplified model.

Time and energy units in DMD

The necessary potential energy associated with hydrogen bond formation in DMD acts as our unit of energy, defined as $E_{HB} = 1.0$. We use this value to express the parameters of the implicit solvent as a unit-less ratio, e.g., E_{HP}/E_{HB} and E_{CH}/E_{HB} for the effective hydropathic and electrostatic interaction strengths, respectively. The temperature of the system is also defined in conjunction with the hydrogen bond energy as $T \sim E_{HB}/k_b$. DMD does not advance each simulation step by a predefined increment of time, rather the time to the next occurring collision between i-j pairs is calculated and used to advance the other molecules in the simulation. The total simulation time and average time of each simulation step can be derived from the equipartition theorem, and at a physiological temperature of 310K ($T = 0.130 \times E_{HB}/k_b$), 10×10^6 time units $\approx 3\mu s$. The physiological temperature of T = 0.130 was determined by Lam *et. al.* in their DMD study on the arctic mutant of Aβ[141], and has been adopted in several studies on Aβ oligomerization[21;70;73;74;142]. Setting the temperature T = 0.130 in turn defines the actual value of the hydrogen bond energy, $E_{HB} = 4.56kcal \times mol^{-1}$. These values are used in the study on αS in Chapter 4.

Implicit solvent: Amino acid-specific hydropathy and electrostatic interactions

The solvent in our system is not explicit, rather an implicit solvent is implemented by the DMD4B-HYDRA force field via hydropathy and electrostatic interactions between side-chains[20;143]. Hydrophobic residues prefer conformations that shield it from the solvent, a behavior mimicked by an attractive force between two hydrophobic residues within a given interaction distance, which defines the boundary of the corresponding potential well. Hydrophilic residues will similarly experience a repulsive force. The strength of the hydrophobic interaction, E_{HP}, together with the Kyte-Doolittle hydropathy scale[3] (Table 1.1), quantifies the hydrophobicity of each amino acid. E_{CH} defines the effective electrostatic interaction between amino acids. E_{HP} has been tuned previously to a value of 0.3 by matching DMD4B-HYDRA-derived oligomer distributions to experimental data[20;21;143]. This value for E_{HP} is independent of the specific system under study, assuming that the experiments which the simulation data is compared to were conducted in an aqueous solution. When choosing an E_{CH} value however, variation in the solvents which are used in experimental studies of different peptides, or even variation in studies of the same peptide, must be accounted for[20;21;143]. Effective hydropathic and electrostatic interactions may individually be positive or negative when applied to specific amino acid interactions, and together can result in a potential well or a potential barrier in the DMD4B-HYDRA force field depending upon the interacting amino acids.

2.3 Multi-Scale Computational Approach

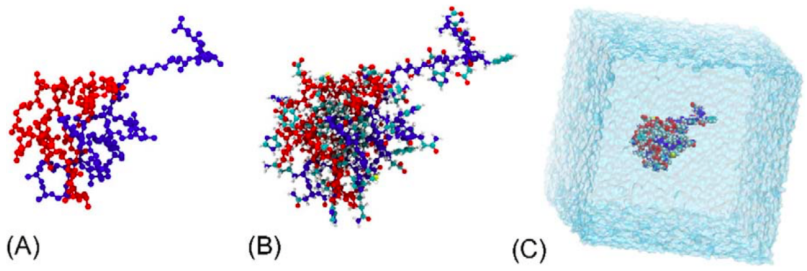

Figure 2.3: Multi-scale computational approach. (A) Four-bead representation of an Aβ40 dimer; (B) Fully atomistic representation of the same Aβ40 dimer; (C) Solvated conformation which is ready for simulation by MD. Image taken from reference[60].

As described in Chapter 3 of this thesis, we took a multi-scale approach in simulations of Aβ40

and Aβ42 monomers and oligomers. DMD4B-HYDRA was first used to simulate the aggregation of unfolded and spatially separated peptides into oligomers. Oligomers present at the conclusion of the DMD4B-HYDRA simulations were then converted into fully-atomistic representations using *prots View*, an in-house software package that replaces the C_β atom of the four-bead conformation with a fully-atomistic side-chain template, followed by energy optimization via the Monte Carlo method. The resulting conformations were then prepared and subjected to explicit-solvent MD. This process can be visualized in Figure 2.3. Explicit-solvent MD cannot reasonably attain the time scales and system size necessary to observe oligomerization from a large system of initially unfolded and dissociated monomers, and DMD4B-HYDRA forfeits a fully-atomistic protein model in order to attain such time scales. By combining these two computational methods, we hoped to overcome some of the inherent shortcomings associated with both MD and DMD4B-HYDRA. This approach allows for the observation of the dynamics involved in oligomer assembly, as well as the folding of fully-atomistic monomers and oligomers in an explicit solvent. This methodology carries the potential for novel insights into the nature of IDPs which may not be observed by studies which use DMD4B-HYDRA or MD alone. The results of this approach are elucidated in Chapter 3 of this thesis.

Chapter 3: Fully-Atomistic Aβ40 and Aβ42 Oligomers in Water: Observation of Porelike Oligomers

This study has been published in the Journal of Chemical Theory and Computation (September 2017)[67]

3.1 Introduction

Alzheimer's disease (AD) is a progressive neurological disorder associated with dementia that is becoming a growing public health issue. As the average lifespan of humans increases due to advances in medicine and technology, it is increasingly pertinent to develop treatments for AD, which affects the growing elderly population. The hallmarks of AD are amyloid plaques, neurofibrillary tangles, and excessive neuronal loss[144]. There is substantial evidence that supports the causative role of soluble, low molecular weight (LMW) oligomers formed by amyloid β-protein (Aβ) in triggering AD[39–45]. Two predominant Aβ alloforms, the 40 and 42 amino acids-long Aβ40 and Aβ42, are normally produced in the brain from amyloid precursor protein by sequential proteolytic cleavages by β- and γ-secretases[35;36]. Despite a relatively small difference in the primary structure between Aβ40 and Aβ42, whereby Aβ42 is characterized by an additional two hydrophobic amino acids at the C-terminus (I41-A42), Aβ42 has an increased propensity to aggregate *in vitro* when compared to Aβ40[37;38]. The hydrophobicity of these amino acids has a profound impact on the distinct role of Aβ42 in AD, whereby increased absolute amounts of Aβ42 or increased Aβ42 to Aβ40 ratio result in increased severity of the disease[145]. Aβ42 has also been shown to be the more toxic of the two alloforms[44], both *in vitro*[146;147] and *in vivo*[148;149].

Oligomers formed by Aβ40 and Aβ42 are difficult to characterize experimentally. Soluble oligomers of various orders coexist with monomers, resulting in a heterogeneous state with a relatively short lifetime. Unlike fibrils, Aβ40 and Aβ42 oligomers lack any ordered structure and have not yet been crystallized, so their three dimensional structure is not known. Bitan *et al.* examined oligomer populations formed by these two peptides via photoinduced cross-linking of unmodified

proteins (PICUP) combined with gel electrophoresis (SDS-PAGE) and showed that the two peptides oligomerize through distinct pathways[150]. Distinct Aβ40 and Aβ42 oligomer populations have been observed also by ion mobility-mass spectrometry (IMS-MS)[151]. Williams *et al.* reported that Aβ40 and Aβ42 oligomers can be cross-linked in the presence of copper and hydrogen peroxide (CHICUP), which leads to oligomer size distributions consistent with previous findings[152]. These experiments showed that Aβ40 oligomer abundance decreases with the oligomer order. In contrast, Aβ42 forms a more complex population of oligomers; the propensity of oligomers decreases from dimers to tetramers, similar to Aβ40, however there is an increased propensity for pentamers and hexamers. Larger Aβ42 oligomers, such as decamers through dodecamers, were also reported in the population of Aβ42 oligomers[150;151].

Computational studies have elucidated several aspects of Aβ monomer folding and oligomer formation[153;154]. Full-length Aβ peptides belong to the class of intrinsically disordered proteins (IDPs), which lack a well-defined native fold. This represents a challenge to computational studies, which need to examine relatively large populations of monomers and oligomers to adequately sample the conformational space and reliably characterize their structure. Despite many advances in computational efficiency, explicit-solvent fully atomistic MD studies of Aβ oligomerization starting from spatially separated and unstructured peptides are still not feasible. To increase sampling efficiency, a multiscale MD approach has been proposed whereby an ensemble of conformations is first acquired through more efficient coarse-grain (CG) modeling, and then the resulting conformations are studied by fully-atomistic MD in explicit solvent[60;155]. One such CG approach combines discrete molecular dynamics (DMD), an efficient event-driven version of MD, and a four-bead peptide model, in which each amino acid is represented by up to four beads, with an implicit solvent force field to account for effective hydropathic and electrostatic interactions (DMD4B-HYDRA)[20]. The DMD4B-HYDRA approach captured distinct Aβ40 and Aβ42 oligomer size distributions[20;21], which match well the experimentally observed distributions[150;151]. Importantly, this approach predicted distinct structural characteristics of Aβ40 and Aβ42 monomers and oligomers[20;21;141] and elucidated the effect of amino acid modifications associated with naturally occurring Aβ variants[21;73;74;156]. DMD4B-

HYDRA simulations predicted that the N-terminal region D1-D7 is more solvent exposed and flexible in Aβ42 oligomers than in Aβ40 oligomers[20;21]. Effective peptide inhibitors of Aβ42 toxicity drastically reduced this N-terminal flexibility and solvent exposure, whereas ineffective inhibitors did not[70]. The solvent accessibility and flexibility of the N-terminal region has thus been implicated in increased toxicity of Aβ42 relative to Aβ40 oligomers.

It is important to test the validity of the DMD4B-HYDRA approach, as the approximations in the amino acid representation and in the force field may impact the observed structural properties of Aβ conformations. In a previous study, large ensembles of Aβ40 and Aβ42 monomers and dimers derived by DMD4B-HYDRA simulations have been examined by MD in explicit water[60]. Albeit somewhat larger and less compact, the resulting fully-atomistic Aβ40 and Aβ42 conformations resembled the DMD4B-HYDRA conformations[60]. Dimer formation was shown to enhance structural differences between the two alloforms, consistent with DMD4B-HYDRA predictions[60]. Fully atomistic Aβ42 dimers were shown to have higher solvent exposure than Aβ40 dimers, as well as increased flexibility at the N-terminal region D1-R5. Here, we expanded the explicit-solvent MD study of monomers and dimers reported by Barz and Urbanc[60] to trimers, tetramers, and pentamers of Aβ40 and Aβ42 and structurally characterized and compared these conformations. By examining how conformations change with oligomer order and elucidating alloform-specific differences in oligomer structure, we furthered our understanding of inherent differences between Aβ40 and Aβ42 and gained insight into the structure-toxicity relationship relevant to AD pathology.

3.2 Results

In the previous study, Barz and Urbanc selected Aβ40 and Aβ42 monomers and dimers obtained by DMD4B-HYDRA simulations, converted them into fully atomistic representation, and used them as initial conformations for MD simulations in explicit water using OPLS-AA force field and two water models, TIP3P and SPCE[60]. Here, we extended this study to include Aβ40 and Aβ42 trimers, tetramers, and pentamers obtained by DMD4B-HYDRA simulations, converted them into fully-atomistic conformations as described previously[60] (additional information is provided in *Methods*), and used them as initial conformations for MD simulations with OPLS-AA force field and TIP3P

water model. We then structurally analyzed all resulting MD trajectories to make a comprehensive comparison of fully atomistic Aβ40 and Aβ42 monomer through pentamer conformations (Table 3.1).

The sequence of Aβ42 is

$$\text{DAEFRHDSGY}_{10}\ \text{EVHHQKLVFF}_{20}\ \text{AEDVGSNKGA}_{30}\ \text{IIGLMVGGVV}_{40}\ \text{IA}_{42},$$

where the first 40 amino acids correspond to the Aβ40 sequence and the C-terminal dipeptide, $\text{I}_{41}\text{A}_{42}$, is present only in A$\beta$42. In the following, we refer to several regions along the sequence: the N-terminal region (D1-D7, 'NTR'), central hydrophobic cluster (L17-A21, 'CHC'), central folding region (A21-A30, 'CFR'), midhydrophobic region (I31-V36, 'MHR'), and the C-terminal region (V39-V40 or V39-A42 for Aβ40 or Aβ42, respectively, 'CTR'). Note that with the exception of the CTR, all other peptide regions are identical for both alloforms.

Table 3.1: Number of MD trajectories for Aβ40 and Aβ42 monomers through pentamers used for analysis. Specific time frames used for analysis are discussed in the *Methods* section.

Species	**Monomers**[a]	**Dimers**[a]	**Trimers**[b]	**Tetramers**[b]	**Pentamers**[b]
Aβ40	45	40	22	16	5
Aβ42	39	39	22	12	4

[a] Simulation time was 50 ns; [b] simulation time was 200 ns.

3.2.1 Convergence of Aβ Monomer and Oligomer Trajectories

In examining oligomer assemblies formed by an IDP such as Aβ, it is important to establish that all conformations under study converge. We thus monitored the root mean square distance (RMSD) of each trajectory as a function of simulation time. Figure S1 shows monomer and dimer trajectories, which converged within 20 ns, as reported previously[60]. Time frames between 20 ns and 50 ns of each Aβ40 and Aβ42 monomer and dimer trajectory were used for structural analysis. Trimer, tetramer, and pentamer trajectories exhibited somewhat slower convergence due to increased system size. As shown in Figure S1, trimer, tetramer, and pentamer trajectories converged within 100 ns, although one of the Aβ40 pentamer trajectories showed a slight yet steady increase in the RMSD value. Time frames between 100 ns and 200 ns of each Aβ40 and Aβ42 trimer, tetramer, and pentamer trajectory

were used for structural analysis.

3.2.2 Multiscale Approach Efficiently Samples Aβ Conformational Space

One of the challenges of studying oligomer formation by MD alone is a relatively long simulation time required for two or more full-length Aβ peptides to start interacting and forming a stable assembly. This challenge can be bypassed by either using advanced sampling techniques, such as replica exchange, or applying a multiscale computational approach, such as the one utilized in the present study[60;154;155]. Here, we asked how efficient is this multiscale approach in sampling the conformational space.

We selected two reaction coordinates: (a) the distance between the Cα atom of the N-terminal D1 and the center of mass of a monomer or oligomer conformation, NT-CM distance, and (b) the sum of solvent accessible surface area (SASA) values over all hydrophobic residues in Aβ (A, M, F, L, V, and I), hydrophobic SASA. Both the NT-CM distance and hydrophobic SASA distributions were previously shown to discriminate between fully atomistic Aβ40 and Aβ42 dimer conformations[60]. Figure 3.1 shows entire trajectories projected onto the conformational space of these two reaction coordinates, where different colors correspond to different trajectories. These plots show that different trajectories overlapped, which is indicative of a well sampled conformational space. Individual trajectories sampled only a limited region of the conformational space that is available to the peptide. The projections of combined, partially overlapped trajectories resulted in a comprehensive and well sampled region in the conformational space, showcasing the effectiveness of this multiscale approach.

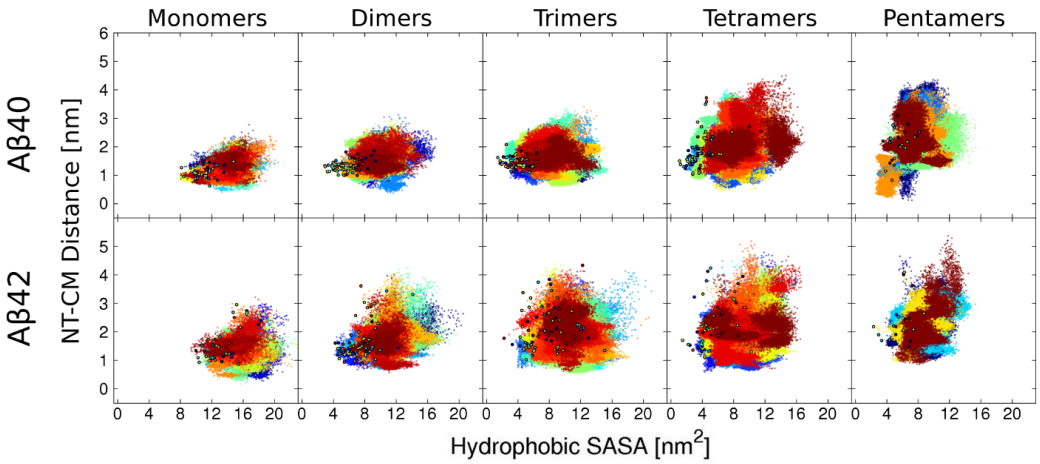

Figure 3.1: Sampling efficiency of Aβ40 and Aβ42 conformations. Characterization of fully atomistic Aβ40 and Aβ42 monomer and oligomer conformations projected onto two reaction coordinates, the NT-CM distance and hydrophobic SASA. Each point in these plots is a projection of a fully atomistic peptide conformation onto the two reaction coordinates. Different colors correspond to different MD trajectories. Black enclosed colored circles correspond to the initial conformations derived by DMD4B-HYDRA. The color of these circles matches the color of the corresponding MD trajectory.

To compare the initial DMD4B-HYDRA conformations with those derived by all-atom MD, Figure 3.1 displays the projections of the initial DMD4B-HYDRA conformations on the two reaction coordinates as colored disks enclosed by black circles. Closer inspection revealed that conformations derived from DMD4B-HYDRA often did not overlap with fully atomistic conformations of the corresponding MD trajectory. However, the ensemble of initial conformations derived by DMD4B-HYDRA mostly overlapped with the ensemble of MD conformations. Relative to the conformational space covered by fully-atomistic MD trajectories, DMD4B-HYDRA conformations were shifted toward lower hydrophobic SASA values, indicating a more efficient shielding of hydrophobic residues from the solvent. We also noted less variability in the NT-CM distance for Aβ40 conformations derived by DMD4B-HYDRA, although the average values were comparable to the average NT-CM distance values observed in the corresponding MD trajectories. Interestingly, the variability in the NT-CM distance for Aβ42 conformations derived by DMD4B-HYDRA was more similar to the variability observed in the corresponding MD trajectories. Overall, the DMD4B-HYDRA conformations appeared more spread out for Aβ42 than for Aβ40 and thus more representative of the overall fully-

atomistic MD structures. Figure 3.2 shows the free energy landscapes of Aβ40 and Aβ42 monomer and oligomer conformations derived from histograms in Figure 3.1.

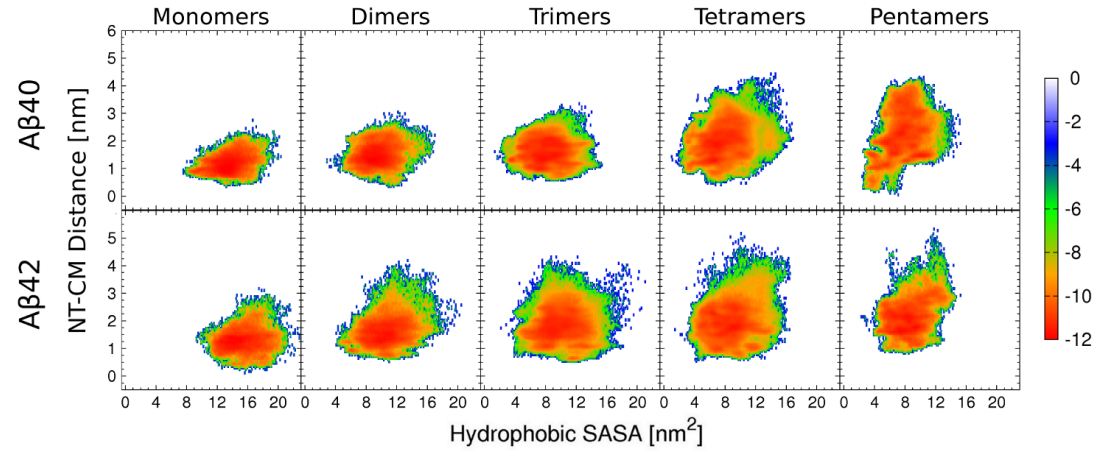

Figure 3.2: Free energy landscapes of Aβ40 and Aβ42 conformations. PMF plots of fully atomistic Aβ40 and Aβ42 monomer through pentamer conformations projected onto two reaction coordinates, the NT-CM distance and hydrophobic SASA. Reaction coordinates were calculated for each individual peptide from each conformation. The color scale on the right is given in units of k_BT.

3.2.3 Trimers Display Largest Alloform-Specific NT-CM Distance and Hydrophobic SASA Differences

To gain more quantitative insight, we calculated one-dimensional (1D) distributions of each of the two reaction coordinates for Aβ40 and Aβ42 monomer and oligomer conformations (Figures 3.3 and 3.4, Table S1). Figure 3.3 (left panel) shows that for monomers, dimers, and trimers, the NT-CM distances in Aβ42 were overall shifted to higher values than in Aβ40, whereby the difference in the average NT-CM distance between Aβ42 and Aβ40 increased with oligomer order, i.e. this difference was the largest for trimers (Table S1). Aβ40 tetramers, and even more so Aβ40 pentamers, showed wider NT-CM distance distributions with larger average values than the corresponding Aβ42 conformations, which is in opposition to DMD4B-HYDRA predictions[20;21].

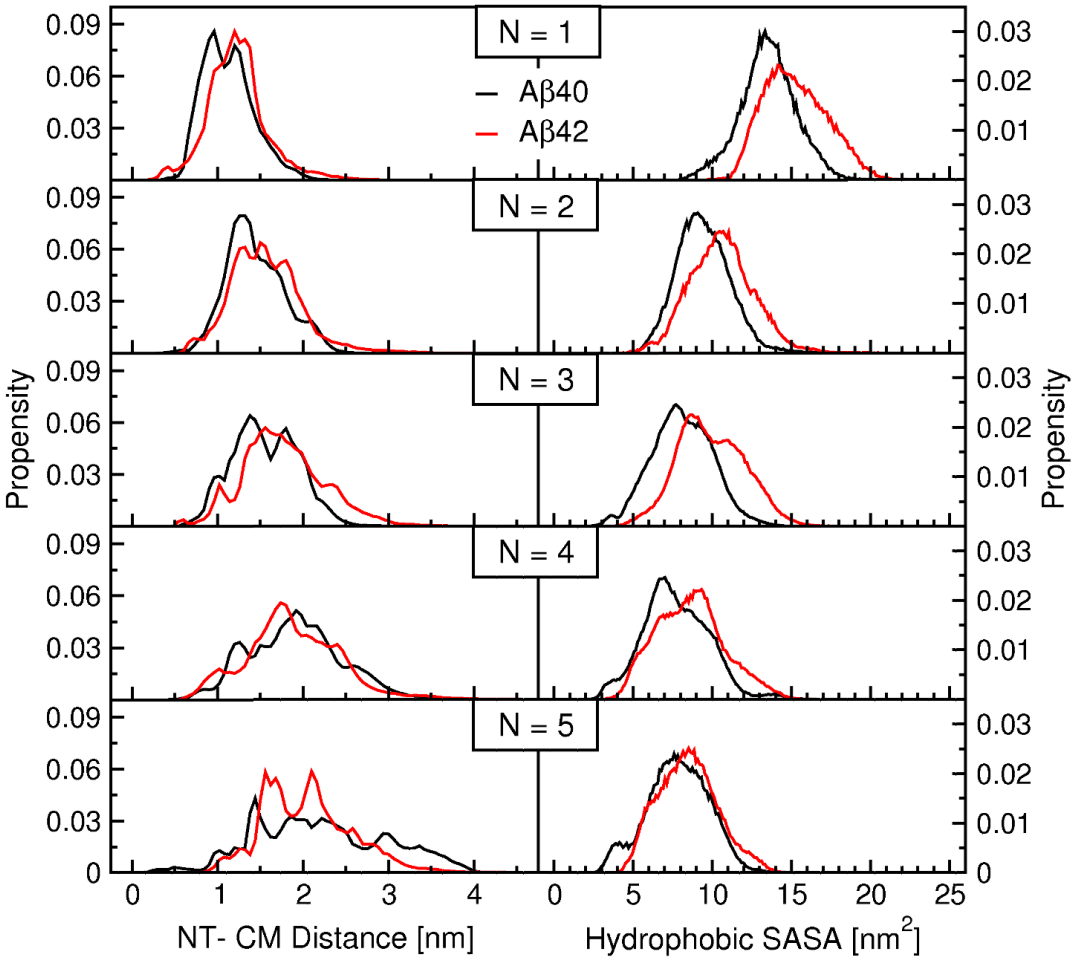

Figure 3.3: Distributions of NT-CM distance and hydrophobic SASA: Aβ40 versus Aβ42. Mann-Whitney U tests comparing Aβ40 and Aβ42 distributions for monomers and oligomers (for both reaction coordinates) resulted in $p < 10^{-3}$ for all.

Distributions of hydrophobic SASA values are displayed in Figure 3.3 (right panel). Notably, Aβ42 hydrophobic SASA distributions were shifted to larger values relative to the respective Aβ40 distributions for monomer through pentamers (Figure 3.3, right panel). As with the NT-CM distance distributions, the largest difference in average hydrophobic SASA values between Aβ40 and Aβ42 was observed for trimers (Table S1). These results demonstrate that of all oligomer sizes, Aβ40 and Aβ42 trimers were the most distinct from each other.

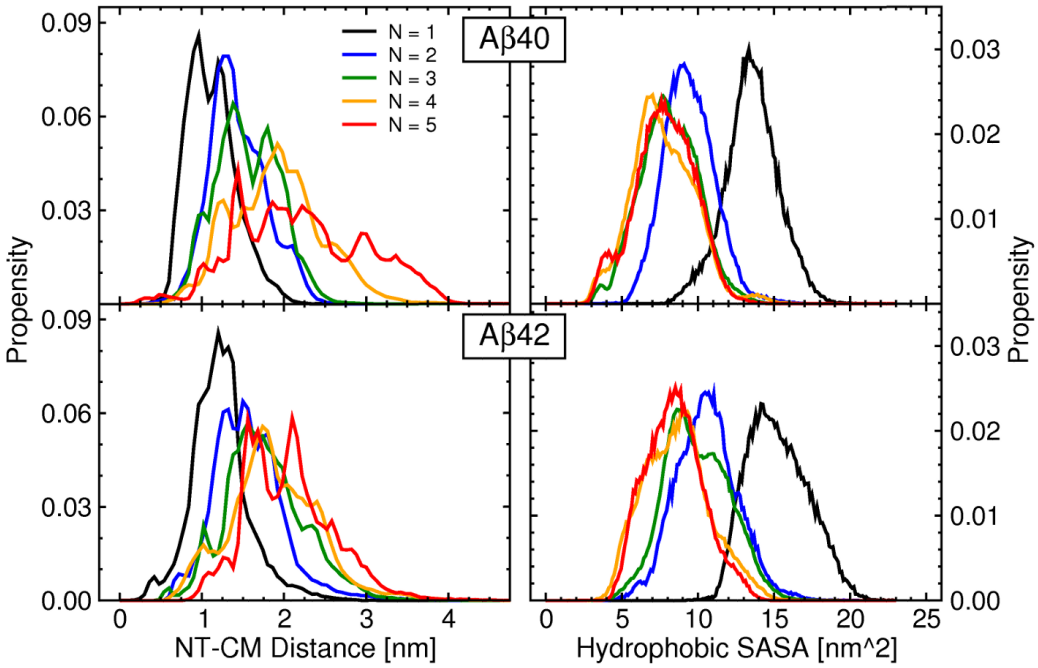

Figure 3.4: Distributions of NT-CM distance and hydrophobic SASA: monomers through pentamers. Mann-Whitney U tests revealed that all distributions for each alloform were distinct from each other (for both reaction coordinates) with $p < 10^{-3}$ for all.

Figure 3.4 shows how the NT-CM distance and hydrophobic SASA distributions changed with oligomer order. The NT-CM distance distribution clearly shifted to larger values and broadened with increased oligomer order for both Aβ40 and Aβ42 (Figure 3.4, left panels). This is in an overall agreement with DMD4B-HYDRA predictions[20;21]. Hydrophobic SASA distributions of Aβ40 and Aβ42 conformations (Figure 3.4, right panels) were shifted to lower values with increased oligomer order, with the exception of Aβ40 pentamers, which were characterized by somewhat larger hydrophobic SASA values than Aβ40 tetramers. Oligomer formation is dominated by a hydrophobic collapse, during which unfavorably solvent exposed hydrophobic residues become shielded from the solvent; this increased hydrophobic SASA in Aβ40 pentamers relative to Aβ40 tetramers may explain why *in vitro* Aβ40 pentamers are less stable (and thus less abundant) than Aβ40 tetramers[150;152].

3.2.4 Aβ40 and Aβ42 Conformations Are Dominated by Statistical Coil and Turn Content

We examined the secondary structure of Aβ40 and Aβ42 monomer and oligomer conformations using STRIDE[157] as implemented in VMD (see *Methods*). We first calculated the average turn, β-strand, bridge, helix, and coil content per peptide in each assembly state (Tables 3.2). The results demonstrate that the secondary structure was dominated by statistical coil ($43 - 49\%$) and turn ($40 - 45\%$), with relatively smaller contributions from β-strand ($4 - 8\%$), bridge ($3 - 5\%$), and α-helical ($< 1.5\%$) content. All statistically significant changes in the average secondary structure of oligomers relative to monomers are marked in Table 3.2 by asterisks. The turn content decreased in Aβ40 and Aβ42 oligomers relative to monomers, but this decrease was only significant in Aβ40 trimers and pentamers and Aβ42 dimers and tetramers, respectively. The β-strand content in Aβ42 conformations was consistently higher than in the respective Aβ40 conformations but differences were not statistically significant. Of all assembly states, Aβ40 and Aβ42 trimers had the highest average β-strand content.

Table 3.2: STRIDE-derived average secondary structure propensities of Aβ40 and Aβ42 conformations. Values are averaged over the entire conformation and error value correspond to SEM values.

Aβ40

N	Turn [%]	β-strand [%]	Bridge [%]	Helix [%]	Coil [%]
1	45.08 ± 1.28	4.24 ± 0.80	4.08 ± 0.38	0.67 ± 0.14	45.90 ± 1.10
2	42.81 ± 1.10	4.67 ± 0.73	3.79 ± 0.32	1.00 ± 0.24	47.71 ± 1.07
3	41.01 ± 1.02*	5.88 ± 0.94	4.02 ± 0.30	1.25 ± 0.24*	47.82 ± 1.45
4	42.63 ± 1.22	5.05 ± 0.66	4.12 ± 0.30	0.93 ± 0.23	47.26 ± 0.83
5	40.71 ± 0.97*	5.62 ± 1.18	4.61 ± 0.91	1.07 ± 0.40	47.98 ± 1.05

Aβ42

N	Turn [%]	β-strand [%]	Bridge [%]	Helix [%]	Coil [%]
1	44.27 ± 1.14	5.90 ± 0.97	3.54 ± 0.36	1.36 ± 0.28	44.92 ± 1.12
2	41.08 ± 1.19*	5.36 ± 0.52	4.06 ± 0.31	0.92 ± 0.17	48.57 ± 0.70*
3	43.42 ± 1.28	7.65 ± 1.05	4.07 ± 0.31	1.12 ± 0.28	43.69 ± 0.91
4	41.02 ± 1.12*	6.17 ± 1.19	4.21 ± 0.25	1.11 ± 0.29	47.48 ± 0.89*
5	43.05 ± 2.10	7.39 ± 1.02	3.88 ± 0.38	1.00 ± 0.22	44.68 ± 1.62

* Values are significantly different than monomers of the same alloform

3.2.5 Distinct Region-Specific Secondary Structure of Aβ40 and Aβ42 Conformations

A detailed description of the secondary structure propensities across specific peptide regions and the corresponding comparison between Aβ40 and Aβ42 conformations is provided in *Appendix A* (see *Detailed Region-Specific Differences in Secondary Structure Between Aβ40 and Aβ42 Conformations*).

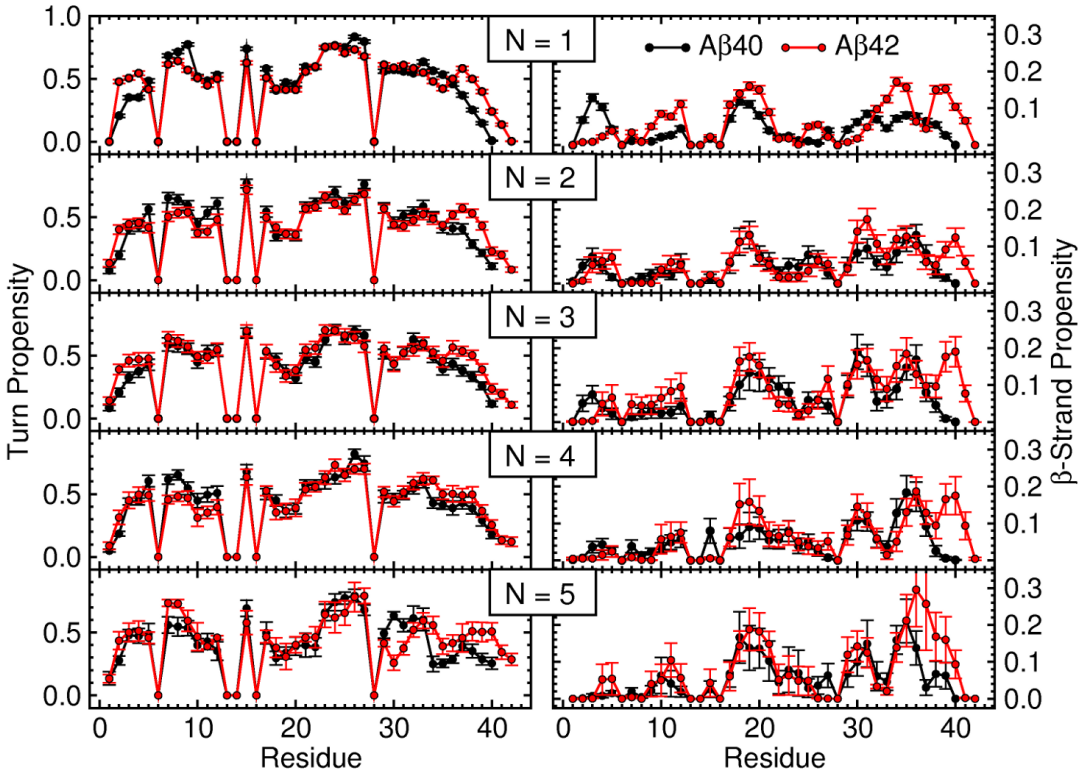

Figure 3.5: Average turn and β-strand (backbone in an extended conformation) propensities per amino acid. Secondary structure was obtained via the STRIDE program in VMD. The error bars are SEM values.

Figure 3.5 displays turn and β-strand propensities for each amino acid in the sequence. In Aβ40 monomers, the highest turn propensities were found at G9, Q15, and in the D23-N27 region. The highest turn propensities in Aβ42 monomers were also noted in the D23-N27 region. This observation for Aβ42 monomers, along with a lower but still significant turn propensity in the G33-G37 region, is in agreement with results of a combined MD/NMR study which reported a prominent turn structure in these regions[158]. High turn propensity in Aβ42 monomers at S26-N27 is also in agreement with NMR data from the study by Ball *at al.*, although the turn propensity observed in our study is three fold higher than the reported experimental value[159]. High turn propensities in the D23-N27 and G33-G37 regions were present in Aβ42 oligomers as well. The most distinct turn propensities between Aβ40 and Aβ42 conformations were observed in the MHR and CTR, wherein Aβ42 conformations had significantly larger turn propensities than Aβ40 conformations

with the exception of tetramers, which showed comparable turn propensities for both alloforms in these regions.

Aβ40 and Aβ42 monomers exhibited the most distinct β-strand propensities of all assembly states. In comparing β-strand content at the C-terminal region, we observed that Aβ42 monomers contained significantly more β-strand structure than Aβ40 monomers within two regions, spanning residues I32-M35 and G38-V40. This is in agreement with previous studies of Aβ40 and Aβ42 monomers, both computational and experimental[158–160]. The largest difference in β-strand propensities between Aβ40 and Aβ42 oligomers was noted at the CTR, where Aβ42 oligomers displayed significantly higher β-strand content than Aβ40 oligomers. DMD4B-HYDRA-derived Aβ40 (but not Aβ42) oligomer conformations displayed high β-strand propensities in the A2-F4 region[21], whereas fully atomistic Aβ40 oligomers displayed increased β-strand propensities in this region only for monomers and trimers.

Coil propensities per amino acid are shown in Figure 3.6 (left panels). In all conformations, regardless of the assembly state and alloform, the coil propensity was at a maximum value of 1.0 at residues H6, H13, H14, K16, and K28. Across all assembly states, Aβ40 conformations displayed significantly larger coil propensities than Aβ42 conformations in the G37-V40 region, which is consistent with the more structured CTR in Aβ42 relative to Aβ40 conformations. Bridge propensities per amino acid are displayed in Figure 3.6 (right panels). Overall, bridge propensities were low and did not exceed $\sim$0.1. Regions with relatively high bridge propensities roughly coincided with hydrophobic regions.

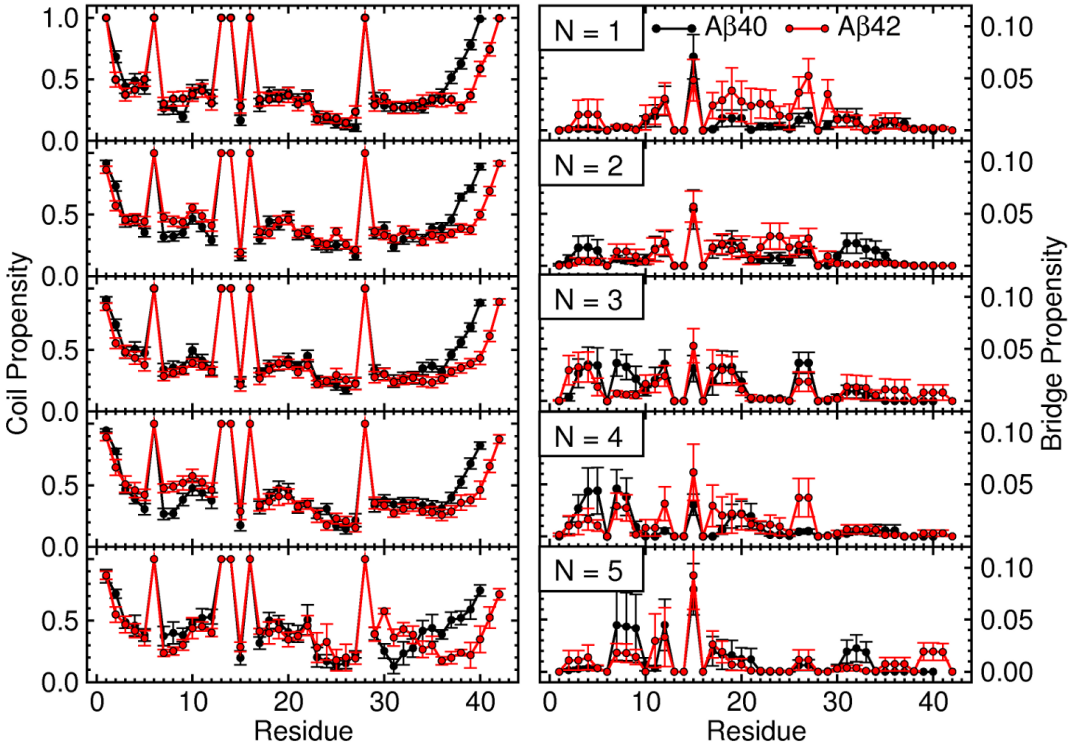

Figure 3.6: Average statistical coil and bridge propensities per amino acid. Secondary structure was obtained via the STRIDE program in VMD. The error bars are SEM values.

3.2.6 Tertiary and Quaternary Structure of Aβ40 and Aβ42 Conformations

To glance into the tertiary and quaternary structure of Aβ40 and Aβ42 conformations, we calculated intra- and intermolecular contact maps as defined in *Methods* for each alloform and each assembly state. Figure 3.7 shows the contacts between pairs of C_α atoms, whereby the lower left triangle and upper right triangle of each contact map correspond to intra- and intermolecular contacts, respectively. The strongest tertiary contacts occurred close to the diagonal and are referred to as diagonal contacts to distinguish them from the more distant, off-diagonal contacts. Note that strong diagonal tertiary contacts contribute to the local stability of turns and loops, whereas strong off-diagonal contacts bring distant peptide regions into proximity. Contacts between pairs of C_β atoms can be found in *Appendix A* (Figure S2).

The predominant off-diagonal tertiary contacts are enclosed in three black frames in Figure 3.7.

Region 1 comprises off-diagonal CHC-CFR, CHC-MHR, and CHC-CTR tertiary contacts (Figure 3.7, black trapezoid-like frames). Region 2 contains off-diagonal tertiary contacts of the MHR with the CTR (Figure 3.7, black triangle-like frames). Region 3 corresponds to tertiary contacts of the NTR with the Q15-V40 region (Figure 3.7, black rectangular frames). To facilitate a comprehensive description of quaternary contacts, we defined three additional regions that were involved the most in quaternary contact formation (Figure 3.7, white frames). Region 4 encloses quaternary CHC-CHC, CHC-MHR, and CHC-CTR contacts (Figure 3.7, white trapezoid-like frames). Region 5 comprises quaternary MHR-MHR, MHR-CTR, and CTR-CTR contacts (Figure 3.7, white triangle-like frames). Region 6 covers quaternary contacts between the NTR and the Q15-V40 region and is thus analogous to Region 3 as defined for tertiary contact maps (Figure 3.7, white rectangular frames). A comprehensive description of contact maps per region is provided in *Appendix A* (see *Detailed Description of Quaternary Structure of Aβ40 and Aβ42*).

Distinct Tertiary Structure of Aβ40 and Aβ42 Monomers

Aβ42 monomers had a larger number of strong diagonal tertiary contacts within the CFR than Aβ40 monomers, implying a more stable turn or loop structure within the CFR of Aβ42 monomers. (Figure 3.7). When examining these same diagonal contacts in the C_β–C_β contact maps, we observed that they were notably weaker (Figure S2). This region is of interest because it was hypothesized to nucleate Aβ folding[161;162], and to adopt a prominent loop in Aβ40 and Aβ42 fibrils[64;163;164]. Previous computational studies of Aβ(21-30) suggested that the turn or loop in this region might be stabilized by a hydrophobic contact between V24 and the butyl group of K28[162;165]. However, the above results indicate that the tertiary contact V24-K28 was dominated by the proximity between the respective C_α rather than C_β groups, suggesting that hydrophobicity was not the prevailing mechanism behind this contact. Of the off-diagonal tertiary contacts found in monomers, those in Region 1 were the strongest for both C_α–C_α (Figure 3.7, black trapezoid-like frames) and C_β–C_β (Figure S2, black trapezoid-like frames) contact maps. Here, the contacts between the CHC and MHR were particularly prominent. Thus, the hydrophobic tertiary contacts between the CHC and MHR, which flank the CFR, most likely stabilize a turn/loop in the CFR.

We also observed strong tertiary contacts in Region 2 between I32-V36 and V39-I41 in both C_α–C_α and C_β–C_β contact maps, which were stronger in Aβ42 than in Aβ40 monomers (Figures 3.7 and S2, black triangle-like frames). This difference in the tertiary structure between Aβ40 and Aβ42 monomers at the C-terminus is consistent with DMD4B-HYDRA predictions[20] and with findings of other experimental and computational studies[158–160;166].

Both C_α–C_α and C_β–C_β tertiary contact maps for Aβ40 and Aβ42 monomers exhibited distinct contacts in Region 3 (Figures 3.7 and S2, black rectangular frames), whereby the contacts involving the NTR were significantly stronger in Aβ40 than in Aβ42 monomers (for a concise description, see the *Tertiary Structure of Aβ40 and Aβ42 Monomers* section of *Appendix A*). A strong involvement of the N-terminus of Aβ40 (but not Aβ42) monomers in tertiary contact formation was predicted by the DMD4B-HYDRA approach[20;21]. Increased intramolecular contacts between the NTR and CHC regions in Aβ40 monomers relative to Aβ42 monomers were also observed in implicit-solvent fully-atomistic replica-exchange MD simulations by Yang and Teplow[160].

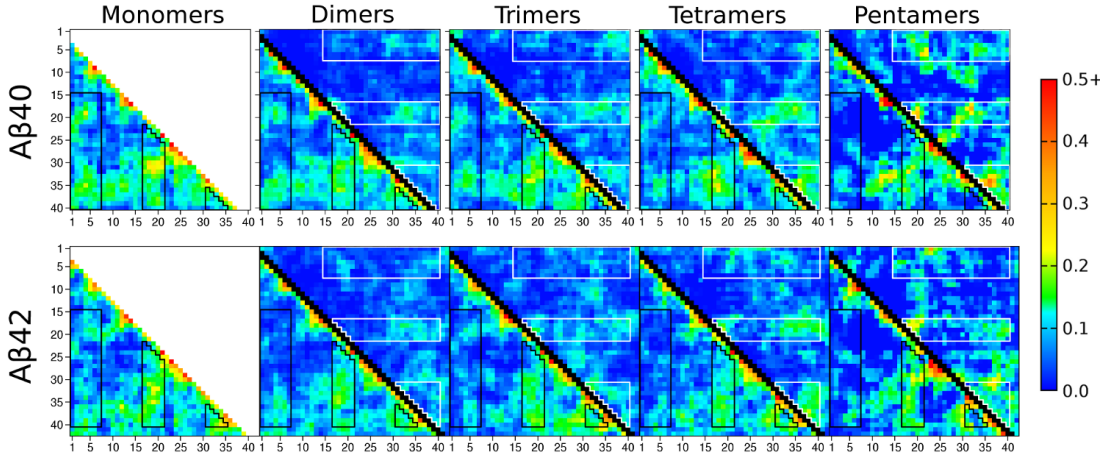

Figure 3.7: C_α–C_α contact maps. Intramolecular contacts are below the black diagonal line, intermolecular contacts are above. The color scale on the right shows the contact propensity as defined in *Methods*. SEM values for all contact propensities were < 0.05 (data not shown).

Involvement of the N-Terminal Region in Tertiary Structure of Oligomers

With an exception of the NTR, oligomers mostly exhibited tertiary structure similar to monomers of their respective alloform (Figure 3.7 and S2). The involvement of the NTR in formation of tertiary contacts in dimers, trimers, and tetramers was also alloform-specific. The NTR of Aβ40 dimers,

trimers, and tetramers was significantly more involved in formation of tertiary C_α–C_α and C_β–C_β contacts than the NTR of the corresponding Aβ42 oligomers (Figures 3.7 and S2, black rectangular frames), consistent with predictions of the DMD4B-HYDRA approach[20;21]. In Aβ40 pentamers, there were fewer tertiary contacts between the NTR and remaining peptide regions than in smaller Aβ40 oligomers. Interestingly, in Aβ40 and Aβ42 pentamers, the NTR formed a comparable number and strength of C_α–C_α and C_β–C_β tertiary contacts with other regions. This result departs from DMD4B-HYDRA predictions, where Aβ42 pentamers and hexamers displayed significantly fewer contacts between the NTR and other peptide regions than both smaller Aβ42 oligomers and Aβ40 pentamers[20].

Tertiary Structure of Aβ40 Versus Aβ42 Pentamers

Pentamers showed the most distinct tertiary structure between the two alloforms. Relative to Aβ42 monomers and smaller oligomers, Aβ42 pentamers were characterized by an increased number of strong diagonal C_α–C_α contacts within the CFR and CTR (Figure 3.7). The corresponding C_β–C_β contacts were significantly weaker (Figure S2). No such increase in C_α–C_α diagonal tertiary contacts was noted for Aβ40 pentamers relative to smaller Aβ40 oligomers. These results imply that the tertiary Aβ42 (but not Aβ40) pentamer structure may be in part stabilized by strong diagonal contacts within the CFR and CTR. In Aβ40 pentamers, strong off-diagonal C_α–C_α and C_β–C_β tertiary contacts were observed in Region 1 (Figures 3.7 and S2, black trapezoid-like frames) between hydrophobic residues flanking the CFR, with several C_β–C_β tertiary contacts noted as being stronger than the corresponding C_α–C_α tertiary contacts. Of all Aβ40 and Aβ42 oligomers, Aβ42 pentamers had the strongest C_α–C_α tertiary contacts in Region 1 (Figure 7, black trapezoid-like frames), concomitant with strong C_β–C_β tertiary contacts (Figure S2, black trapezoid-like frames). These contacts enclosed the CFR, which is the region that was hypothesized to nucleate Aβ folding, and represents an essential "loop" feature of the Aβ fibrillar structure[163;167]. These results imply that in Aβ40 pentamers, and even more so in Aβ42 pentamers, the loop within the CFR is stabilized by hydrophobic contacts between the flanking CHC and MHR.

Distinct Quaternary Structure of Aβ40 and Aβ42 Oligomers

Quaternary contacts in Region 4 (Figures 3.7 and S2, trapezoid-like white frames) highlighted the involvement of the CHC in oligomer formation. These quaternary contacts overall increased in either strength or number (or both) with oligomer order in both Aβ40 and Aβ42 oligomers. In Aβ40 and Aβ42 pentamers, quaternary contacts between pairs of CHCs were stronger than in smaller oligomers. Within Region 4 and between Regions 4 and 5, we observed a complex of anti-parallel C_α–C_α quaternary contacts in Aβ40 tetramers that were even more pronounced in Aβ40 pentamers (Figure 3.7, trapezoid-like white frames). Notably, this complex of anti-parallel quaternary contacts was absent from Aβ42 tetramers and pentamers. These results suggest that Aβ40 oligomers might be more prone to form an anti-parallel structure than Aβ42 oligomers.

The strength of C_α–C_α and C_β–C_β quaternary contacts in Region 5, which encloses contacts between MHRs, between CTRs, and between the MHR and CTR, increased with oligomer order in both Aβ40 and Aβ42 oligomers (Figures 3.7 and S2, white triangle-like frames). Thus, in addition to the CHC, the MHR and CTR were also involved in oligomer formation. In Aβ42 pentamers, the strongest C_α–C_α quaternary contacts were partially shifted out of Region 5 toward the CFR. Overall, the C_α–C_α and C_β–C_β quaternary contacts in Region 5 were stronger in Aβ40 than in Aβ42 pentamers, in contrast to DMD4B-HYDRA predictions, which revealed significantly more quaternary contacts between CTR regions in Aβ42 than in Aβ40 oligomers[20;69]. Aβ40 tetramers and pentamers were also characterized by diagonal quaternary contacts between MHRs and CTRs. Of all Aβ40 and Aβ42 oligomers, the diagonal quaternary contacts within and adjacent to Region 5 spanned the widest region in Aβ42 pentamers (Figures 3.7 and S2, white triangle-like frames).

The involvement of the NTR in the quaternary structure was specific to oligomer order (Region 6 in Figures 3.7 and S2, white rectangular frames). In Aβ40 dimers and trimers, the NTR was more involved in C_α–C_α and C_β–C_β quaternary contact formation than in Aβ42 dimers and trimers, in agreement with DMD4B-HYDRA predictions[20;21]. In contrast, Aβ40 tetramers exhibited fewer quaternary contacts in Region 6 than Aβ42 tetramers. Interestingly, both Aβ40 and Aβ42 pentamers were characterized by strong yet distinct quaternary contacts in this region. Whereas NTR-CHC

and NTR-MHR contacts dominated the quaternary structure of Aβ40 pentamers in Region 6, in Aβ42 pentamers the strongest C_α–C_α and C_β–C_β quaternary contacts in Region 6 were centered around D1-D23.

Peculiar Role of D1 in Tertiary and Quaternary Structure

In monomers, D1 formed tertiary contacts with E22, D23, and the C-terminal amino acid V40 (Aβ40) or A42 (Aβ42). These contacts were facilitated by effective electrostatic interactions between the positively charged amino group of D1 and the negatively charged side chain groups of E22 and D23 or the negatively charged carboxyl group of V40 (Aβ40) or A42 (Aβ42). These contacts were somewhat stronger in Aβ40 than in Aβ42 monomers (Figure 3.7 and S2, black rectangular frames). Aβ40 dimers were still characterized by tertiary contacts D1-E22 and D1-V40, however the contact D1-D23 was significantly weaker (Figure 3.7 and S2, black rectangular frames). In Aβ40 trimers and tetramers, the tertiary D1-E22 contact was still present, whereas the tertiary D1-V40 contact was much weaker. In Aβ40 pentamers, D1 did not form strong tertiary contacts. In Aβ42 dimers and larger oligomers, tertiary contacts D1-E22, D1-D23, and D1-A42 were diminished relative to Aβ42 monomers.

Quaternary contacts involving D1 in dimers were not significant. Three quaternary contacts, D1-D1, D1-E3, and D1-E11 were noted in Aβ40 trimers and only one strong quaternary contact, D1-A42, was present in Aβ42 trimers. Tetramers displayed two quaternary contacts in each alloform, D1-D1 and D1-V40 in Aβ40 and D1-E22 and D1-K28 in Aβ42. Aβ40 pentamers were characterized by a strong quaternary contact, D1-E22. Interestingly, in Aβ42 pentamers, D1 formed one of the strongest quaternary contacts with D23, and to a lesser extent also with E22. Moreover, each Aβ42 pentamer trajectory was characterized by this strong quaternary contact.

We further explored the origin of this D1-D23 quaternary contact in Aβ42 pentamers. Aspartic acid is unique among amino acids because it can form a hydrogen bond between its side chain and its backbone amino group, leading to an increased propensity to adopt turn-like conformations[168]. Because each aspartic acid can simultaneously act as a hydrogen bond donor and acceptor in an intraresidue hydrogen bond formation, two aspartic acids could potentially form two interresidue

hydrogen bonds by pairing up their backbone amino groups with their C_γ oxygens. We extracted all distances between: (i) the two C_γ oxygens of D1 and the amino group of D23 and (ii) the amino group of D1 and the two C_γ oxygens of D23, and plotted the two respective histograms (Figure 3.8, panels A and B). Comparison of the two histograms clearly shows distinct features at small distances. The histogram of distances (i) displayed a clear peak at $5 - 10$ Å (Figure 3.8A). However, only the histogram of distances (ii) revealed a high propensity for distances < 5 Å with a peak at ~ 2.5 Å, consistent with hydrogen bonding (Figure 3.8B). The existence of a hydrogen bond between the amino group of D1 and the side chain of D23 was confirmed through hydrogen bond analysis within VMD[169]. This asymmetry in hydrogen bond formation is a consequence of the positively charged backbone amino group of D1 (but not D23). The peak in the distance histogram (i) at $5 - 10$ Å was simply a consequence of the proximity between D1 and D23. Inspection of the arrangement of water molecules around the quaternary contact D1-D23 in Aβ42 pentamers revealed that water molecules were often absent from the D1-D23 contact region (Figure 3.8C), although water bridges between the functional backbone groups of D1 and D23 were sometimes noted (data not shown). The above results demonstrate that the positively charged N-terminus (D1) played a significant role in the quaternary structure of Aβ42 pentamers.

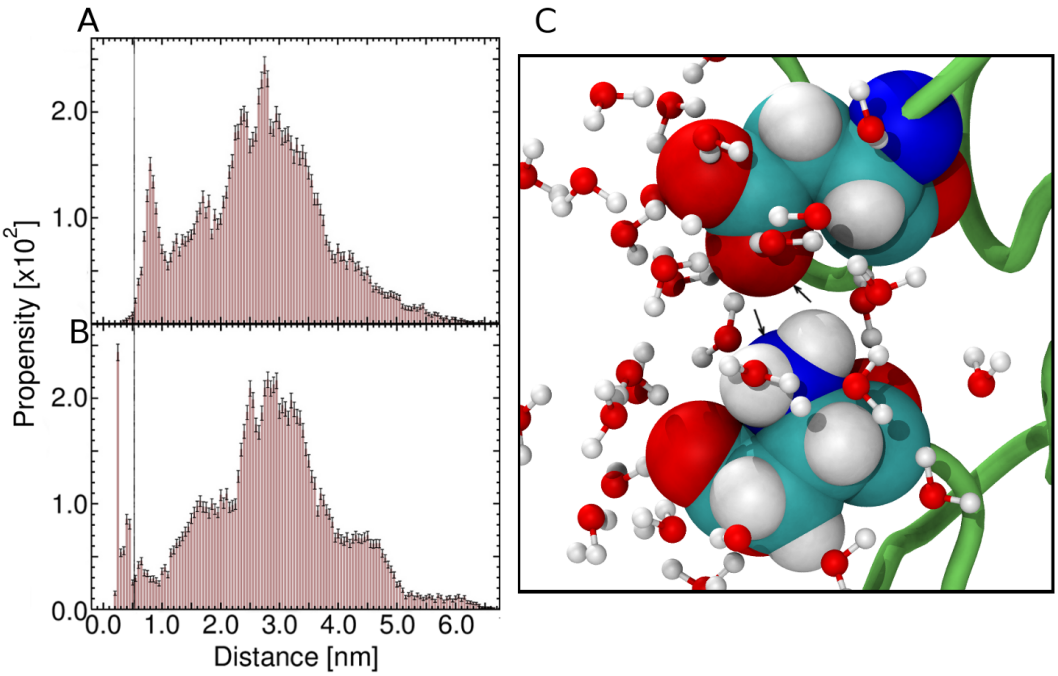

Figure 3.8: Elucidation of the quaternary D1-D23 contact in Aβ42 pentamers. Normalized histograms of the distance between (A) the C_γ carboxyl oxygens of D1 and the backbone amino group of D23 and (B) the backbone amino group of D1 and the C_γ carboxyl oxygens of D23, derived from all Aβ42 pentamer trajectories (time frames $100 - 200$ ns). The black line placed at 5.25 Å is a visual guide. The error bars for (A) and (B) are SEM values. (C) Snapshot of the quaternary contact between the backbone N atom of D1 (blue sphere with a black arrow) and the C_γ carboxyl oxygen of D23 (red sphere with a black arrow) from an Aβ42 pentamer trajectory. Water molecules surrounding the two residues, D1 and D23, are also depicted.

3.2.7 Amino Acid Arrangement within Oligomers

We contrasted the arrangement of individual amino acids in the two alloforms by calculating the average distance of each residue from the center of mass (CM) for Aβ40 and Aβ42 monomers through pentamers (Figure S3, left panel). Aβ42 monomers, dimers, and trimers were characterized by larger distances from the CM in the N-terminal A2-R5 region than the corresponding Aβ40 conformations. Differences between the two alloforms in the NTR were the largest for trimers, with Aβ42 trimers being farther from the CM than Aβ40 trimers, in agreement with DMD4B-HYDRA predictions[20;21;170]. The largest difference between Aβ40 and Aβ42 conformations was observed at the CTR, where Aβ42 was associated with significantly smaller distances from the CM than Aβ40. A more detailed comparison is provided in *Appendix A* (see *Detailed Distance from the Center of*

Mass (CM) Analysis).

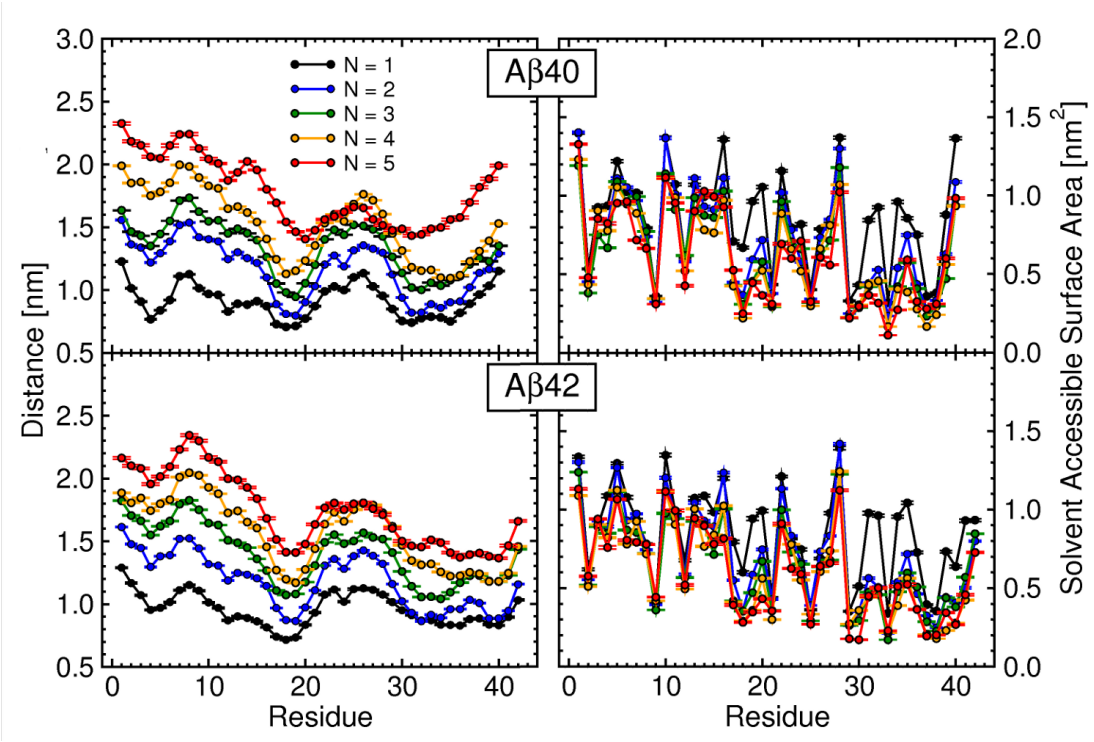

Figure 3.9: Distance from the CM and SASA per amino acid: assembly state effect. The error bars are SEM values.

We also plotted the distance from the CM per residue for each alloform individually to elucidate how the amino acid arrangement with respect to the CM changed with oligomer order (Figure 3.9, left panels). Overall, distances increased with oligomer order. In all Aβ40 conformations (except for trimers) and in Aβ42 monomers, dimers, and trimers, D1 was on average the most distant from the CM. In Aβ40 trimers, Aβ42 tetramers, and Aβ42 pentamers, S8 assumed the largest distance from the CM. Overall, the NTR was the farthest region from the CM for all oligomers, and the distance from the CM of the NTR with respect to other regions increased with oligomer order.

3.2.8 Alloform- and Assembly State-Specific Solvent Exposure

Solvent accessibility of different peptide regions represents an important aspect of oligomer structure, as the most solvent exposed amino acids would have an increased propensity to interact with cellular targets. We thus calculated the SASA per amino acid for each Aβ40 and Aβ42 assembly state (Figure S3, right panel). A larger number of hydrophobic amino acids were exposed to the solvent

in Aβ42 monomers, dimers, and trimers than in Aβ40 monomers, dimers, and trimers (6 versus 4, 5 versus 2, and 5 versus 1). V40 was consistently more solvent accessible in Aβ40 than in Aβ42 conformations. This result was not unexpected because the sequences of Aβ40 and Aβ42 differ at the CTR and the negative charge of the C-terminus should expose the CTR to the solvent in an alloform-specific way. Interestingly, A2 was more accessible to the solvent in Aβ42 than in Aβ40 conformations, regardless of the assembly state. F4 was more solvent accessible in Aβ42 monomers, dimers, and trimers than in respective Aβ40 conformations. Similarly, R5 showed higher solvent exposure in Aβ42 than in Aβ40 monomers, dimers, tetramers, and pentamers. These results are consistent with the reported increased SASA at the NTR in Aβ42 relative to Aβ40 oligomers[20;170].

We also asked to what extent the SASA per residue would be affected by the assembly state for each individual alloform. Figure 3.9 shows the SASA per amino acid for monomers through pentamers of Aβ40 (upper right panel) and Aβ42 (lower right panel). Consistent with an observation of the hydrophobic SASA distributions (Figure 3.4, right panels), which were shifted to lower values with increasing oligomer order (with an exception of Aβ40 pentamers), the SASA per amino acid overall decreased with oligomer order for both Aβ40 and Aβ42. The decrease in SASA values was particularly large between monomers and dimers, whereby the largest decrease in SASA values was observed within regions H14-E22 and K28-V40 in Aβ40 and L17-F20 and G29-A42 in Aβ42, thus mostly coinciding with hydrophobic regions. Because the first 10 amino acids in the sequence were at the largest distance from the CM in both Aβ40 and Aβ42 oligomers, we closely examined the SASA per amino acid in this region as described in *Appendix A* (see *Detailed Alloform- and Assembly State-Specific Solvent Exposure Analysis*). An overall decrease of SASA values in the NTR upon oligomer formation indicated that the NTR was involved in contact formation, in agreement with an increased frequency of the corresponding quaternary contacts in the NTR for Aβ40 and Aβ42 tetramers and pentamers (Figure 3.7).

3.2.9 Salt Bridge Formation

At neutral pH, the sequence of Aβ40 (and Aβ42) contains three positively charged (R5, K16, and K28) and six negatively charged (D1, E3, D7, E11, E22, and D23) amino acids. Salt bridges that can

form between two oppositely charged amino acids can sometimes contribute to the structural stability of proteins or protein complexes. We thus calculated salt bridge propensities for all positively (Figure S4, left panel) and negatively (Figure S4, right panel) charged amino acids in the Aβ sequence.

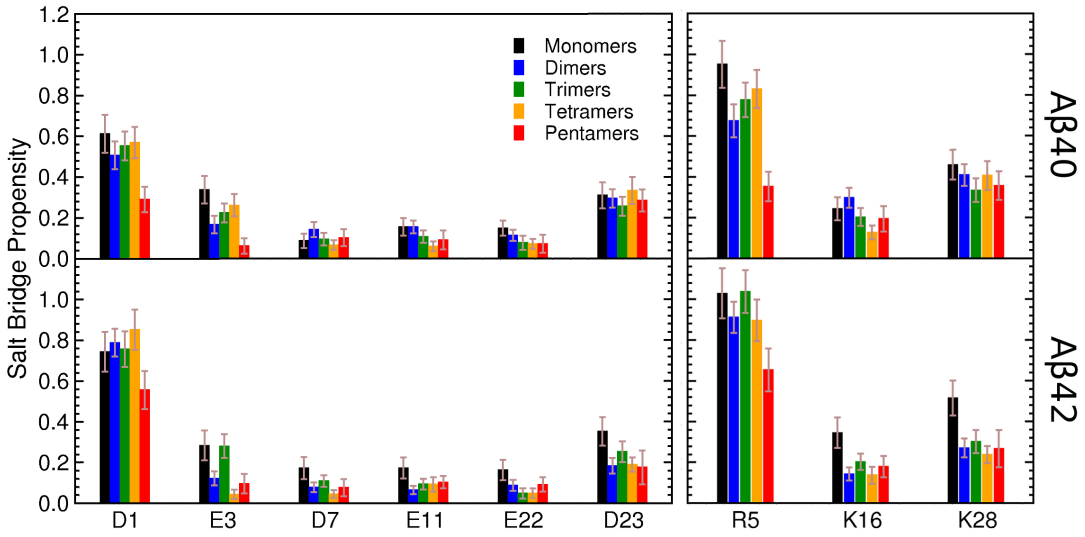

Figure 3.10: Histograms of salt-bridge propensities: assembly state effect. Sums of intra and interpeptide propensities, for the three positive (right panel) and six negative (left panel) amino acids are shown. The error bars are SEM values.

Negatively Charged Amino Acids

Among the six negatively charged amino acids, D1 was the most involved in salt bridge formation in both Aβ40 and Aβ42 conformations, followed by D23 and E3. The remaining negatively charged amino acids D7, E11, and E22 showed the lowest salt bridge propensities (Figure S4, left panel). Aβ40 and Aβ42 monomers showed comparable D1 salt bridge propensities whereas Aβ42 oligomers showed significantly larger D1 salt bridge propensities than Aβ40 oligomers. Aβ40 tetramers and pentamers displayed higher D23 salt bridge propensities than the corresponding Aβ42 conformations. In addition, Aβ40 tetramers had a significantly higher E3 salt bridge propensity than Aβ42 tetramers. D1 salt bridge propensities were comparable among Aβ40 monomers through tetramers, however, Aβ40 pentamers displayed a significantly lower D1 salt bridge propensity (Figure 3.10, left panel). A similar trend in D1 salt bridge propensities was noted in Aβ42 conformations, where

Aβ42 pentamers had a significantly lower D1 salt bridge propensity than smaller Aβ42 oligomers (Figure 3.10, left panel). D23 salt bridge propensities were comparable in all Aβ40 conformations. Aβ42 dimers, tetramers, and pentamers had lower D23 salt bridge propensities than Aβ42 monomers (Figure 3.10, left panel). Dimers and pentamers of both Aβ40 and Aβ42 showed decreased E3 salt bridge propensities relative to monomers, and the same trend was observed in Aβ42 tetramers (Figure 3.10, left panel). In Aβ40, D7, E11, and E22 salt bridge propensities did not strongly depend on the assembly state, whereas in Aβ42, these propensities decreased in oligomers relative to monomers (Figure 3.10, left panel).

Positively Charged Amino Acids

Of the positively charged amino acids, R5 was the most involved in salt bridge formation across all assembly states in both Aβ40 and Aβ42, though Aβ40 pentamers were characterized by comparable R5 and K28 salt bridge propensities (Figure S4, right panel). Aβ40 dimers and tetramers showed higher K28 salt bridge propensities than Aβ42 dimers and tetramers. Aβ42 dimers and pentamers were characterized by higher R5 salt bridge propensities than the corresponding Aβ40 conformations (Figure S4, right panel). There were no significant alloform-specific differences in K16 salt bridge propensities. Pentamers exhibited lower R5 salt bridge propensities than monomers (Figure 3.10, right panel). K16 salt bridge propensities were comparable for all Aβ40 conformations except for tetramers, which exhibited a decreased propensity. In all conformations K28 showed higher salt bridge propensities than K16, but the difference was statistically significant only in Aβ40 monomers, trimers, tetramers, and pentamers. Aβ42 oligomers had lower K16 and K28 salt bridge propensities than Aβ42 monomers (Figure 3.10, right panel).

Do Salt Bridges Contribute To Aβ Oligomer Formation?

Tables S2-S6, which contain more detailed information about intra and interpeptide salt bridge propensities for specific pairs of amino acids, revealed further differences between the two alloforms and illuminated the dominance of intrapeptide over interpeptide salt bridge occurrences. Salt bridge propensities that showed statistically significant differences between the two alloforms are shown in red font in Tables S2-S6. Overall, the highest propensities were associated with intrapeptide

D1-R5 and E3-R5 salt bridges. D1-R5 had a significantly higher propensity to form intrapeptide salt bridges in Aβ42 than in Aβ40 oligomers, and occurred in 52% of Aβ42 tetramer conformations (Table S5). The intrapeptide E3-R5 salt bridge occurred in over 20% of all conformations. The D23-K28 intrapeptide salt bridge was more likely to occur in Aβ40 tetramers and pentamers than in the corresponding Aβ42 conformations. The highest interpeptide salt bridge propensity of 6% was observed between R5-E11 for Aβ40 dimers and R5-E22 for Aβ42 dimers. The remaining interpeptide salt bridge propensities occurred in $< 6\%$ of all conformations and could thus not have contributed considerably to oligomer structure.

3.2.10 Morphology of Fully Atomistic Aβ Conformations

In order to extract representative Aβ40 and Aβ42 oligomer conformations, we recalculated the free energy landscapes shown in Figure 3.2 by assigning a single, average value to each oligomer conformation in the reaction coordinate space. While the free energy landscapes for monomers were not affected, the resulting free energy landscapes for oligomers shown in Figure 3.11 populated a smaller region of phase space than the original landscapes (Figure 3.2), but allowed us to identify oligomer conformations with the lowest free energy. All free energy landscapes were characterized by broad minima, within which the free energy variations were small compared to thermal energy (k_BT). For each assembly state of each alloform, we selected 250–1300 ($\sim$3.5%) representative conformations with the lowest free energy. Figure 3.11 displays several representative Aβ40 and Aβ42 monomers and oligomer conformations associated with the free energy minima.

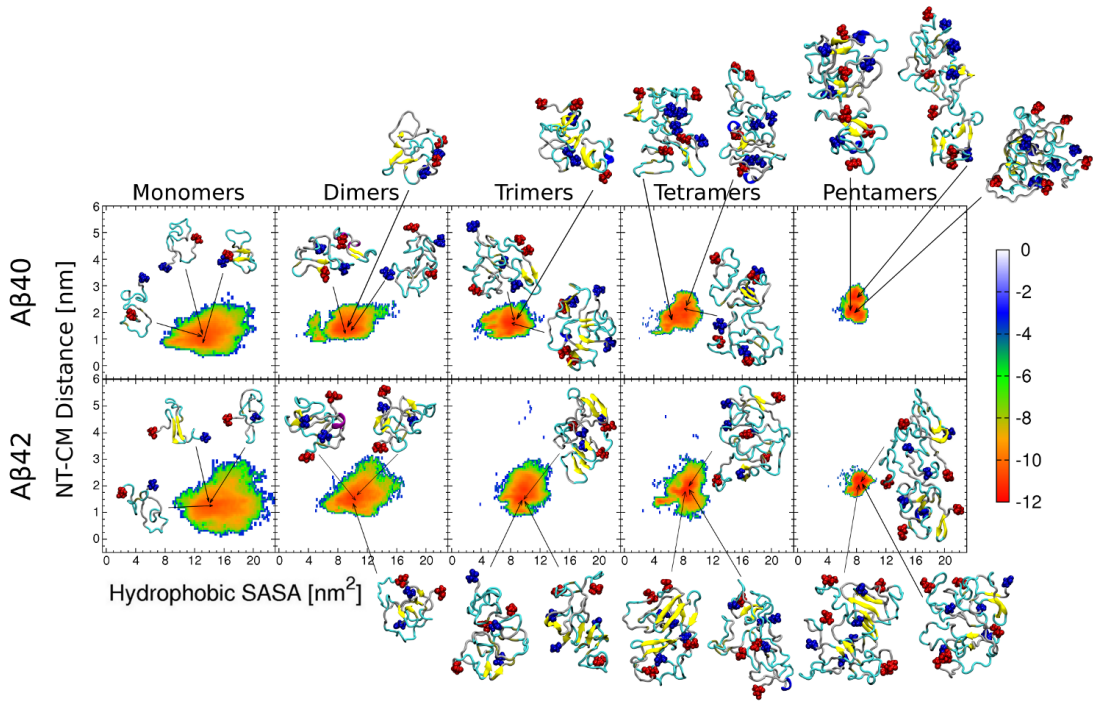

Figure 3.11: Free energy landscapes with representative conformations of Aβ40 and Aβ42 conformations. PMF landscapes of fully atomistic Aβ40 and Aβ42 monomer through pentamer conformations, projected onto two reaction coordinates, the NT-CM distance and hydrophobic SASA, are shown. Reaction coordinates were averaged over all peptides in each oligomer. The color scale on the right is given in units of k_BT.

Representative monomer, dimer, and trimer conformations of both alloforms exhibited overall globular morphologies, which resembled the DMD4B-HYDRA conformations but were larger, less compact, and had a more solvent-exposed CTR. In Aβ42 monomers, dimers, and trimers, the NTR was more exposed to the solvent than the CTR, which made Aβ42 conformations visibly distinct from Aβ40 conformations. In comparison to the DMD4B-HYDRA conformations, fully atomistic Aβ40 and Aβ42 tetramers displayed more elongated morphologies. Initially, only 4 of 16 Aβ40 and 1 of 12 Aβ42 tetramer conformations derived from DMD4B-HYDRA simulations had an elongated morphology. In 9 of 16 all-atom MD trajectories, Aβ40 tetramers adopted a distinct elongated morphology which, once formed, did not change with simulation time. A close inspection of each trajectory revealed that the Aβ40 tetramer morphology was stabilized either within the first 50 ns (in 9 of 16 trajectories), or at a significantly longer time between 100 and 150 ns (in 7 of 16 trajectories). Relative to Aβ40 tetramer conformations, Aβ42 tetramer conformations mostly

retained their globular morphologies; elongated Aβ42 tetramer conformations were observed in only 3 of 12 trajectories. An inspection of Aβ42 tetramer trajectories revealed that the morphology was always stabilized by ∼100 ns. Aβ42 tetramers were more prone to exhibit elongated conformations than Aβ42 trimers, although this difference was notably smaller than the corresponding difference between Aβ40 trimers and Aβ40 tetramers. A similar tendency to elongate was also observed for Aβ40 pentamers, which were predominantly elongated (in 4 out of 5 trajectories) and displayed distinct anti-parallel β-strands, whereas Aβ42 pentamers displayed a mix of globular, elongated, and disk-shaped morphologies.

3.2.11 Observation of porelike Oligomer Conformations

A prominent hypothesis about the mechanism of Aβ oligomer-induced toxicity states that Aβ oligomers form pores in a cellular membrane that act as ion channels, causing an abnormal flux of ions into the cell[105;106;110]. Initial Aβ40 and Aβ42 conformations derived from DMD4B-HYDRA simulations did not display pore–like morphologies. We here examined all fully atomistic Aβ40 and Aβ42 trajectories to assess the ability of oligomers to form annular porelike structures. Aβ40 and Aβ42 monomers and dimers did not show any porelike morphologies. In contrast, numerous Aβ40 and Aβ42 trimer, tetramer, and pentamer conformations were characterized by at least one pore with a diameter >1 Å. Most of these conformations with small pores formed intermittently and were not particularly long lived. For most of them, the pore diameter did not exceed 2 Å, which would be too small to allow water molecules or Ca^{2+} ions (with a diameter of ∼2.3 Å) to pass through. Of the 22 Aβ40 and 22 Aβ42 trimer conformations, 2 and 3 conformations, respectively, revealed pores with diameters ranging from 3–5 Å. Among the 16 Aβ40 and 12 Aβ42 tetramer conformations, 4 and 3, respectively, contained pores with diameters ranging from 3–5 Å (in Aβ40 tetramers) or 3–6 Å (in Aβ42 tetramers). Of 5 Aβ40 pentamer trajectories, 3 showed pore–like morphologies with a pore diameter of 3–6 Å. porelike morphologies observed in Aβ40 pentamers, as well as those seen in trimers and tetramers of both alloforms, contained pores with depths of 5–10 Å. All 4 Aβ42 pentamer trajectories displayed conformations with pores, with pore diameters ranging from 3–7 Å and depths of 6–12 Å, which exceeded the depth of pores formed by other oligomers.

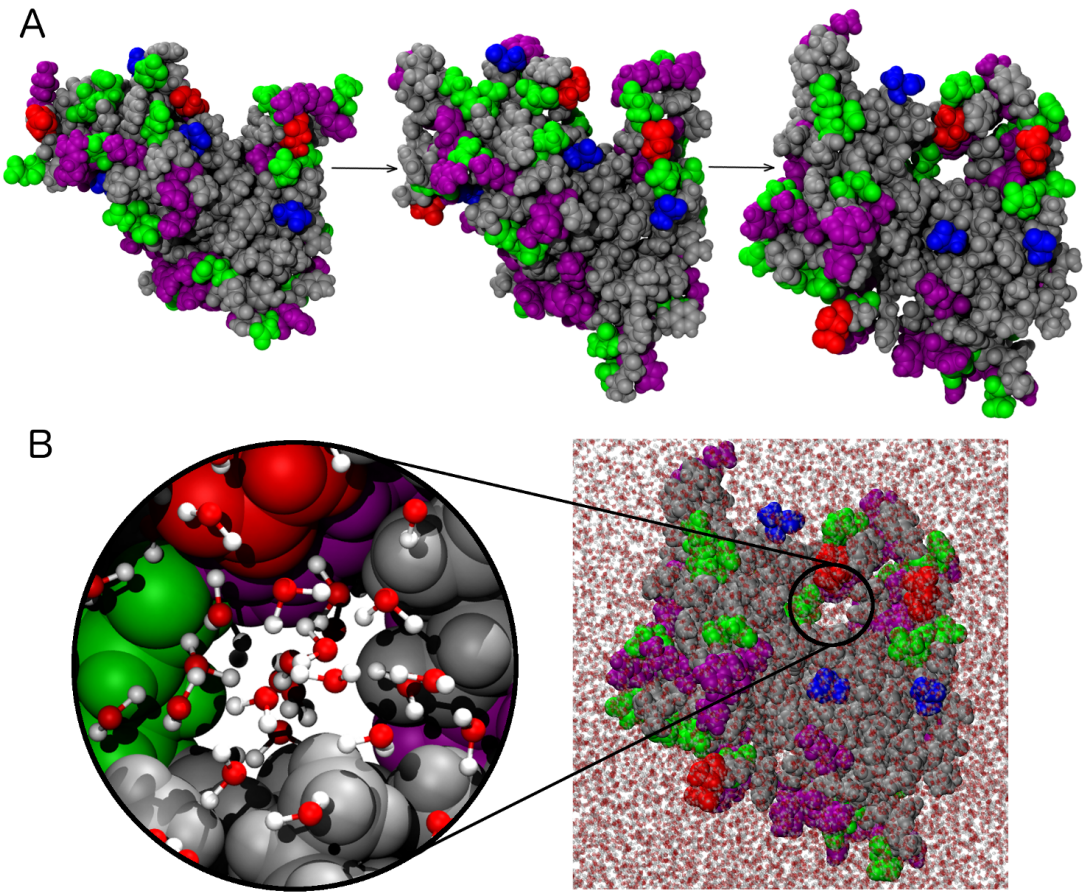

Figure 3.12: Aβ42 pentamer with a porelike morphology. (A) Snapshots of a pore formation within an Aβ42 pentamer trajectory at 5 ns, 20 ns, and 200 ns. (B) The Aβ42 pentamer conformation with the pore immersed in water (water oxygen and hydrogens are shown as red and white spheres) as observed at 200 ns. A magnified pore structure is shown on the left. The N-terminal and C-terminal amino acids D1 and A42 are shown as red and blue spheres, respectively. The negatively and positively charged amino acids are displayed as green and purple spheres, respectively.

Figure 3.12A shows a progression of pore formation in one of the Aβ42 pentamer trajectories, and Figure 3.12B shows the same pentamer in the presence of water molecules to demonstrate that the pore was large enough to allow water to flow through (see insert on the left). A closer inspection of this Aβ42 pentamer trajectory revealed that the pore formed within the first 25 ns and remained stable throughout the the remainder of the 200 ns-long trajectory. As observed in Figure 3.12A, where the N-terminal amino acid D1 is shown as a red sphere, the N-terminus on the left approached another N-terminus on the right, closing the structure into a pore. In addition to

D1, three other negatively charged residues (E3, E22, and D23) were found to reside on the inner surface of the pore. We also identified the positively charged amino acids R5 and K16 in the inner lining of the pore. Apart from R5, K16, and the positively charged backbone amino group of D1, no other positively charged amino acid was observed on the inner pore surface, rendering the inner pore overall negatively charged. If this pentamer with a pore were to be embedded into a membrane, the presence of a negatively charged pore lining could facilitate an abnormal flux of positively charged ions, such as calcium ions. The abnormal influx of calcium ions was proposed to disrupt the calcium homeostasis within neurons in the context of the ion channel hypothesis of Aβ oligomer-induced toxicity[105].

Porelike Aβ oligomers with low β-strand and α-helical content observed here are distinct from previously reported *ad hoc* β-barrel models of Aβ ion channels, i.e. β-strand-rich theoretical constructs made of Aβ peptides, which have been simulated by MD[59;109;119]. In these models, H13 and H14 are positioned in the lining of the ion channel and are hypothesized to play a key role in the inhibition of calcium ion flux via zinc ions[171–173]. We examined the composition of the pore lining in all Aβ oligomers that contained pores and found H6 and H14, but not H13, in the lining of a single porelike Aβ42 pentamer conformation, indicating that histidines did not significantly contribute to the pore lining. It is important to note, however, that in our study porelike Aβ oligomers formed in water, which differs from the hydrophobic environment inside a membrane. Aβ oligomer structure is thus expected to change upon insertion into a membrane. Our observation of Aβ oligomers forming pores in water prior to their interaction with the membrane provides important structural insights into the pathway leading to Aβ oligomer toxicity.

3.3 Discussion

Here, fully atomistic conformations of Aβ40 and Aβ42 monomers through pentamers are examined by MD in explicit water using a multiscale approach, whereby the initial conformations are derived from implicit-solvent DMD4B-HYDRA simulations[21]. As shown in the previous MD study on Aβ40 and Aβ42 monomers and dimers[60] and in the present work, this multiscale approach overcomes two types of challenges: (i) long simulation times that would be otherwise required to simulate oligomer

formation by explicit solvent MD alone and (ii) increased sampling of the phase space required to study assemblies formed by IDPs such as Aβ. The effective sampling of the phase space is achieved by acquiring multiple trajectories from distinct initial conformations. To the best of our knowledge, this is the first comprehensive fully atomistic explicit-solvent study of Aβ40 and Aβ42 monomers and LMW oligomers up to and including pentamers.

Exploration of the free energy landscapes of fully atomistic Aβ40 and Aβ42 conformations revealed alloform-specific differences in the distributions of the two reaction coordinates, the NT-CM distance and hydrophobic SASA. Aβ42 monomers, dimers, and trimers showed significantly larger NT-CM distances and hydrophobic SASA values than the corresponding Aβ40 conformations. These alloform-specific differences increased with oligomer order, such that Aβ40 and Aβ42 trimers were the most distinct. This trend did not continue beyond trimers, although Aβ40 and Aβ42 distributions of NT-CM distance and hydrophobic SASA values were also alloform-specific for tetramers and pentamers. The secondary structure analysis revealed that Aβ40 and Aβ42 conformations were dominated by statistical coil and turn, whereas β-strand, bridge, and α-helical content was significantly lower.

We observed tertiary structure differences between Aβ40 and Aβ42 monomers that mostly affected the CTR in a manner consistent with previous studies[20;141;158–160;166;174;175]. Our results also show that the NTR was significantly more involved in formation of tertiary contacts in Aβ40 monomers, dimers, trimers, and tetramers than in the corresponding Aβ42 conformations, consistent with DMD4B-HYDRA predictions[20]. As expected, hydrophobic regions were the most strongly involved in quaternary contact formation, whereby the contribution of the MHR and CTR increased with oligomer. Of all oligomers, Aβ40 and Aβ42 pentamers were characterized by the most distinct tertiary and quaternary structures. In Aβ42 pentamers, the strongest quaternary contact was formed between the positively charged backbone amino group of D1 and the negatively charged side chain of D23, highlighting the importance of the positively charged N-terminus in the Aβ42 pentamer structure.

Our analysis of the distance from the CM of individual amino acids and solvent accessibility along

the sequence demonstrated that just like in monomers, the CTR remained significantly more exposed to the solvent in Aβ40 than in Aβ42 oligomers. The amino acids in the NTR were overall at the largest distance from the CM and the increase of the distance from the CM in the NTR with oligomer order was larger than the corresponding increase in other regions. As expected, the decrease in SASA values upon oligomer formation was the largest in the three hydrophobic regions (CHC, MHR, and CTR), showing that hydrophobic collapse was driving oligomer formation. The NTR was the most involved in salt bridge formation. In oligomers, the salt bridges with the highest propensity were intramolecular E3-R5 and D1-R5, whereas intermolecular salt bridges did not contribute significantly to oligomer structure. Aβ42 oligomers overall displayed a higher propensity to form salt bridges involving the NTR higher than that involving Aβ40 oligomers.

Finally, we examined the morphology of representative Aβ40 and Aβ42 conformations. As noted already by Barz and Urbanc for monomers and dimers[60], fully atomistic conformations were somewhat larger compared to the initial DMD4B-HYDRA conformations, reflecting the simplified side chain in the latter. Aβ40 and Aβ42 monomers, dimers, and trimers adopted predominantly globular morphologies, in which the N-termini were exposed to the solvent more than the other peptide regions. Aβ40 tetramers showed a distinct tendency to elongate, more so than Aβ42 tetramers. This tendency to elongate was even stronger in Aβ40 pentamers, which is in agreement with DMD4B-HYDRA predictions of dumbbell-shaped Aβ40 pentamers and hexamers[21]. However, the tendency to form elongated morphologies was stronger in fully atomistic than in the corresponding DMD4B-HYDRA conformations. Fully atomistic Aβ42 pentamers showed the most diverse set of morphologies, including globular, elongated (prolate spheroid-like), and disk-shaped (oblate spheroid-like).

The most striking result with direct implications to AD pathology was the observation of porelike morphologies in Aβ40 and Aβ42 trimers, tetramers, and pentamers in pure water. We focused on porelike conformations with a pore diameter larger than 3 Å, which is large enough to allow water to pass through. All observed pores were at least 5 Å in length. Importantly, the occurrence frequency of such porelike oligomer conformations increased with oligomer order and was somewhat higher for Aβ42 than for Aβ40 oligomers. porelike conformations with the deepest pores (6–12 Å in length)

were observed in Aβ42 pentamer trajectories. According to our structural analysis (Figure 3.9, left panels), Aβ42 pentamer conformations were 40–50 Å in diameter, thus approximately matching the thickness of a cellular membrane. We examined one of these pores and noted that its inner surface was overall negatively charged. Thus, if such a porelike pentamer were embedded into a cellular membrane, the negatively charged inner pore lining would promote the flow of positive ions such as calcium through the pore, and thus act as an ion channel. This scenario is consistent with several *in vitro* studies, which reported that in a lipid bilayer Aβ forms calcium ion channels that cause an abnormal calcium flux into the cell, leading to cell death[117;118]. Importantly, our results demonstrate that Aβ trimers, tetramers, and pentamers can adopt porelike conformations prior to their insertion into a membrane.

No MD study to date has reported a porelike morphology of Aβ oligomers formed in pure water. Strödel *et al.* conducted MD simulations of Aβ42 monomers through tetramers, hexamers, and octamers in a membrane, whereby oligomers were constructed from either monomeric or dimeric units and were high in β-sheet content[176]. Two MD studies, Jang *et al.* and Connelly *et al.*, assembled U-shaped Aβ42 monomers into β-barrel structures containing 12-20 Aβ42 peptides, and tested their stability within a membrane by MD[120;121]. MD studies of Aβ pores within a membrane and Aβ oligomer insertion into a membrane[63] typically use oligomers with high β-strand or β-sheet content, whereas the porelike Aβ40 and Aβ42 oligomers observed here are dominated by statistical coil and turn with relatively low β-strand and α-helical content. Typically, theoretical models of ion channels or pores are characterized by a significant amount of ordered structure and possess cylindrical symmetry. It is important to note, however, that although Aβ ion channel models were often constructed to be in agreement with AFM data[121;177], they were not formed from initially unstructured and separated peptides. This is understandable because explicit-solvent MD of protein assembly in a membrane environment is computationally demanding. Insertion of multiple antimicrobial peptides into a membrane and their subsequent formation of ion channels (pores) was examined using MD by Marrink and collaborators[178–180]. They found that the pore structure was notably more disordered than the structure predicted by β-barrel models and was consistent with

experimental data. On the basis of the results of these studies combined with our findings, we anticipate a similarly disordered ion channel structure to be formed by Aβ in a membrane.

One of the key questions in AD pathology is whether or not oligomer toxicity depends on oligomer order. The results of our study suggest that the ability of Aβ oligomers to form pores, and thus the resulting toxicity, sharply increases with oligomer order. This offers a plausible explanation for the reported disproportionately large increase in toxicity of covalently cross-linked Aβ40 oligomers with oligomer order[181]. However, oligomer toxicity depends not only on the propensity for pore formation but also on the relative abundance of oligomers of a specific order, which is expected to decrease with oligomer order both *in vitro* and *in vivo*. Because in our study trimers were the smallest oligomers associated with pore formation, trimers could represent the key oligomer species that mediates toxicity in AD. Indeed, naturally secreted Aβ trimers were shown be more potent inhibitors of hippocampal synaptic plasticity than Aβ dimers both *in vitro* and *in vivo*[182]. Moreover, Sherman *et al.* showed that AD brain-extracted Aβ trimers, but not Aβ monomers and dimers, induced pathological changes in tau protein, which mediate toxicity in AD[183]. On the other hand, *in vitro* Aβ42 (but not Aβ40) forms abundant pentamers and hexamers as well as their multiples, e.g. dodecamers[150;151;184]. Our results show that Aβ42 pentamers have an order of magnitude higher propensity for pore formation ($\sim$100%) than Aβ42 trimers ($\sim$10%). Furthermore, the pore diameters observed in Aβ42 trimers were overall smaller in diameter than the pores observed in Aβ42 penatmers, and the hydration shell around Ca^{2+} ions could impede their passage through the more narrow pores observed in trimers[185]. Thus, Aβ42 pentamers (and potentially hexamers) could represent the proximate toxic species in AD as long as their abundance exceeded $\sim 1/10^{th}$ of the abundance of Aβ42 trimers, which is certainly true for *in vitro* conditions[152;186]. Our findings thus provide structural insights into Aβ oligomer toxicity and support the ion channel hypothesis of AD, which states that Aβ toxicity stems from the ability of Aβ oligomers to form ion channels[98;99;105;106;108].

3.4 Methods

3.4.1 Simulation protocol

The initial conformations for both species were derived via the DMD approach with the four-bead protein model and implicit solvent residue-specific interactions[21]. These conformations were converted into the all-atom representations using an in-house software package, *protsView*. This package uses all-atom side-chain templates to replace the C_β atom of the four-bead conformation, followed by optimization of the contact energy using the Monte Carlo method. Hydrogens were added to the resulting conformations using the VMD software package[169].

Molecular dynamics simulations were then conducted using GROMACS 4.9.5[123–126] with the OPLS force field[129;130], hydrated via the TIP3P[127] explicit water model. Each conformation was placed in a hydrated cubic box extending 30 Å from the surface of the protein and subjected to periodic boundary conditions. NA^+ ions were added to neutralize the net charge of the Aβ peptides in water. All N-termini were positively charged (NH^3+) and C-termini were negatively charged (COO-). Each Aβ-water system was subjected to energy minimization using the steepest descent algorithm, followed by a 200 ps equilibration run with the heavy atoms being constrained to their positions to allow the water molecules to equilibrate around the peptide/oligomer structure. The initial conformations can be seen in the plot in Figure 3.2 as a colored circle with a thick black border. The temperature was held at 310 K by a velocity rescaling thermostat with a stochastic term[187]; a time constant of 0.1 ps was used and the atmospheric pressure was enforced by the Parrinello-Rahman method using a coupling constant of 2 ps[188]. The number of trajectories used in analysis is reported in Table 3.1. Monomers and dimers were simulated for 50 ns. Trimers, tetramers, and pentamers were simulated for 200 ns. Conformations were recorded 50 ps apart, resulting in 1,000 conformations for each monomer and dimer trajectory, and 4,000 conformations for each trimer, tetramer, and pentamer trajectory. The *Structural Analysis* section details which conformations were used for the different types of analysis. The simulations were conducted on *Steele* at Purdue University and *Stampede* at Texas Advanced Computing Center through the NSF TeraGrid and the Extreme Science and Engineering Digital Environment (XSEDE) supercomputing resources.

3.4.2 Structural Analysis

Structural analysis of oligomer species entailed examining each peptide in a given oligomer individually, unless otherwise noted in the manuscript. For example, for each pentamer conformation we examined, a data point is produced for each of the five peptides which compose the pentamer. Average values of per-oligomer analysis is averaged over each peptide within a given oligomer, rather than being averaged over the entire oligomer. Conducting analysis in this manner allows us to examine the structure of peptides in the context of oligomer conformations.

Potential of the Mean Force

The potential of the mean force (PMF) calculation entailed projecting each conformation onto two selected reaction coordinates; the sum of the solvent accessible surface area of the hydrophobic amino acids in the sequence (according to the Kyte-Doolittle scale[3]) and the distance from the N-terminus of each peptide to the center of mass of each monomer/oligomer. Within the conformational space of these two reaction coordinates, we created a normalized histogram and counted the number of conformations, N_i, within each bin of the histogram. These histograms were used to create the distributions for these two reaction coordinates. The PMF values were then calculated via the equation $-k_bT \ln(N_i/N_t)$, where N_t is the total number of conformations for a given assembly state and alloform. Every conformation from every analyzed trajectory was used in the PMF calculation.

Secondary Structure Analysis

Analysis for average secondary structure propensities was calculated as follows: the average secondary structure propensity for each trajectory of a given oligomer size and species is first calculated, both per-residue and averaged over the each peptide individually, using the STRIDE program[157] implemented within the VMD software package[169] to identify secondary structures. These values were then used to calculate a histogram for each secondary structure, whose mean is presented as the average secondary structure propensities in Table 3.2. The error associated with these averages is the standard error of the mean (SEM) for these histograms, calculated as the corrected sample standard deviation divided by the number of trajectories used for analysis. Monomer and dimer tra-

jectories were analyzed using all conformations from the final 30 ns of simulation. Larger oligomers utilized all conformations from the final 100 ns of simulation.

Contact Maps

Two amino acids were considered in contact if the distance between their C_α atoms or C_β- atoms were within a distance of 7.5 Å from each other, consistent with previous DMD and MD studies[20;21;60;69]. The contact maps are presented as an (i,j) matrix with the number of contacts between two specific amino acids averaged over all analyzed conformations. Intramolecular contacts include contact between every (i,j) pair of amino acids within a given peptide (tertiary contacts), and intermolecular contacts include every (i,j) pair of amino acids that belong to different peptides within an oligomer (quaternary contacts). The intramolecular and intermolecular contact maps are presented in a single plot, with intramolecular contacts below the diagonal and intermolecular contacts above the diagonal. Intramolecular contact between an amino acid and any of the two neighboring amino acids on each side are not considered, as their geometry results in these (i,j) pairs falling within the cutoff distance by default. The color scale corresponds to the propensity for an (i,j) pair of amino acids to be in contact, calculated as the number of observed contacts between an (i,j) pair divided by the number of possible contacts that can occur across the analyzed frames. SEM values were not included as they were negligible when compared to the calculated value due the large number of analyzed conformations. Monomer and dimer trajectories were analyzed using all conformations from the final 30 ns of simulation. Larger oligomers utilized all conformations from the final 100 ns of simulation.

Solvent Accessible Surface Area

Solvent accessible surface area (SASA) was calculated within the VMD software package[169]. The calculation utilized a spherical surface 1.4 Å away from the atom's van der Waals surface as the cutoff. This calculation was applied to all atoms in an amino acid to give a resulting SASA value. The SASA-per-residue values were attained by averaging over every analyzed conformation. Error bars correspond to SEM values. Monomer and dimer trajectories were analyzed using all conformations from the final 30 ns of simulation. Larger oligomers utilized all conformations from the final 100 ns

of simulation.

Distance to Center of Mass

The distance to center of mass for each residue was calculated by first determining the center of mass for each monomer/oligomer via the VMD software package[169]. The location of the C_α atom for each residue is calculated and the distance from this atom to the center of mass of the monomer/oligomer is calculated. The data presented is averaged over every analyzed conformation. Error bars correspond to SEM values. Monomer and dimer trajectories were analyzed using all conformations from the final 30 nanoseconds of simulation. Larger oligomers utilized all conformations from the final 100 ns of simulation.

Salt Bridge Propensities

Salt bridge propensities between positively charged amino acids R5, K16 and K28, and negatively charged amino acids D1, E3, D7, E11, E22 and D23 residues were calculated within the VMD software package[169]. The condition for identifying salt bridge formation requires any of the the side-chain nitrogen atoms of the positively charged amino acids to be within 3.2 Å of the side-chain oxygen atom of a negatively charged amino acid. The percent of analyzed frames containing a salt bridge involving a given amino acid or between two given amino acids defines the represented propensity values. Error bars and values correspond to SEM values. Monomer and dimer trajectories were analyzed using all conformations from the final 30 nanoseconds of simulation. Larger oligomers utilized all conformations from the final 100 ns of simulation.

3.5 Acknowledgments

This research was supported in part by the National Science Foundation through XSEDE resources provided by the Texas Advanced Computing Center and Purdue University under the grant number and PHYS100030 (B.U). We thank Dr. Marisa Roman for preparing the Aβ40 and Aβ42 trimer, tetramer and pentamer conformations for analysis and Dr. Luis Cruz for help with the in-house protsView package.

Chapter 4: Folding and Oligomer Formation of α-Synuclein: A Discrete Molecular Dynamics Study

4.1 Introduction

Parkinson's Disease (PD) is a neurodegenerative disorder which impairs motor function by affecting neurons in the substantia nigra region of the brain. PD is characterized not only by neuronal loss, but by the presence of Lewy bodies, abnormal intracellular protein aggregates. The 140 amino acids-long, $\sim$14 kDa α-synuclein (αS), a protein which is expressed in neurons and found in presynaptic nerve terminals, was found to be the principal component of Lewy bodies[29;189;190]. Strong evidence has implicated αS in playing a central role in the pathogenesis of PD, such as familial mutations[77;78;82] or increased expression[79–81] of αS which result in PD pathology[29]. The physiological function of αS is not well understood, although there is evidence that it acts as a chaperone protein and it may play a role in regulating neurotransmitter release and vesicle fusion[191–194]. Although αS exists in Lewy bodies in aggregated amyloid fibrillar form, substantial research, both *in vitro* and *in vivo*, points to intermediate oligomeric species of αS as the primary cytotoxic agents[83–88].

αS belongs to the class of proteins known as intrinsically disordered proteins (IDPs). As such, it lacks a well-ordered and folded native structure, and has been posited to exist in the brain predominantly as an unfolded monomer[195;196]. Alternatively, αS has also been shown to fold into a well-ordered tetramer that has a characteristic α-helical structure and resists aggregation[96]. Further research proposes that αS exists physiologically as a combination of unfolded monomers and a dynamic mixture of oligomeric assemblies[197;198]. Photo-induced cross-linking of unmodified proteins (PICUP) combined with gel electrophoresis (SDS- PAGE) has been used to characterize populations of αS *in vitro*, and found it to be a heterogeneous mixture of monomers and LMW oligomers[199–204]. These experiments largely report that monomers are the most abundant species, followed by low molecular weight (LMW) oligomers up to at least tetramers with a monotonically decreasing propensity. Although, this experimental method is biased towards LMW oligomers as the resolution de-

creases with increasing molecular weight, it provides quantitative measure of the oligomeric species in a soluble state prior to fibril formation.

Experimental characterization of monomer and LMW oligomers of αS is challenging both experimentally and computationally. This has motivated the use of computational approaches to understanding αS, although the size and intrinsically disordered nature of αS makes it difficult to ensure efficient sampling of the conformational phase space. Attempts to overcome this limitation include MD combined with NMR measurements, which impose NMR-derived distance restraints, and implementation of replica-exchange combined with MD[205;206]. DMD combined with a four-bead amino acid model and implicit solvent with amino acid-specific interactions (DMD4B-HYDRA) has been applied previously to simulations of folding and aggregation of IDPs. This computational approach was able to successfully capture oligomer distributions of Stefin B and WT and pathogenic mutations of Aβ40 and Aβ42 which were consistent experimental results [20;21;73;138;150;151]. Furthermore, DMD4B-HYDRA was used to elucidate and predict structural differences between WT Aβ40 and Aβ42 monomers and oligomers[20;21;141], as well as the impact of naturally occurring amino acid substitutions in Aβ which have been associated with early-onset AD[21;73;74;156]. To date, no study has examined soluble αS populations by computational methods. We here apply efficient DMD combined with a four-bead model and implicit-solvent force field in order to study aggregation of αS from initially unfolded and separated peptides. This approach allows for simulation of longer time scales than MD while still providing insights into the structural characteristics of αS monomers and oligomers.

4.2 Results

The primary structure of αS is given in Figure 4.1. We have divided the sequence into three regions; the N-terminal region (NTR, residues M1-K60), the non-Aβ component (NAC, residues E61-V95), and the C-terminal region (CTR, residues K96-A140). The NTR has been implicated in the mediation of binding to phospholipid vesicles[29]. The name NAC is derived from the non-Aβ component of amyloid plaques,which represent one of the hallmarks of AD. The NAC is the most hydrophobic region of the sequence and plays a role in the formation and structure of αS fibrils[29].

Figure 4.1: The 140 amino-acids long αS sequence. Hydrophobic amino acids are displayed in red font, uncharged polar amino acids are shown in green font, charged polar amino acids are shown in blue, and the NAC region is underlined.

The CTR, which is the most hydrophilic of the three regions, is believed to be involved in the function of αS as a chaperone protein[29;92;191].

4.2.1 Tuning the Effective Electrostatic Interaction Strength to Experimental Conditions

In order to find the E_{CH} value which results in the best match to experimental data, we plotted the oligomer size populations at the final time step of our simulations for each E_{CH} value. Figure 4.2 shows the oligomer size distribution resulting from simulations of the five lowest E_{CH} values, ranging from $E_{CH} = 0.00 - 0.30$. The value $E_{CH} = 0.30$ was the highest E_{CH} value for which monomers and LMW oligomers of different orders coexisted after 80 M time units. Simulations with $E_{CH} = 0.40$ and 0.50 produced a single oligomer composed of all 16 αS peptides at the final time step (data not shown). As the E_{CH} value is increased from 0.00 to 0.10, we observed a monotonically decreasing monomer population with a monotonic increase in both dimer and trimer populations. The oligomer population resulting from simulations with $E_{CH} = 0.20$ resulted in decreased both monomer and dimer populations relative to lower E_{CH} values, and correspondingly increased populations of larger oligomers up to 13-mers. E_{CH} values larger than 0.20 produced a distribution which shifted abruptly to higher-order oligomers.

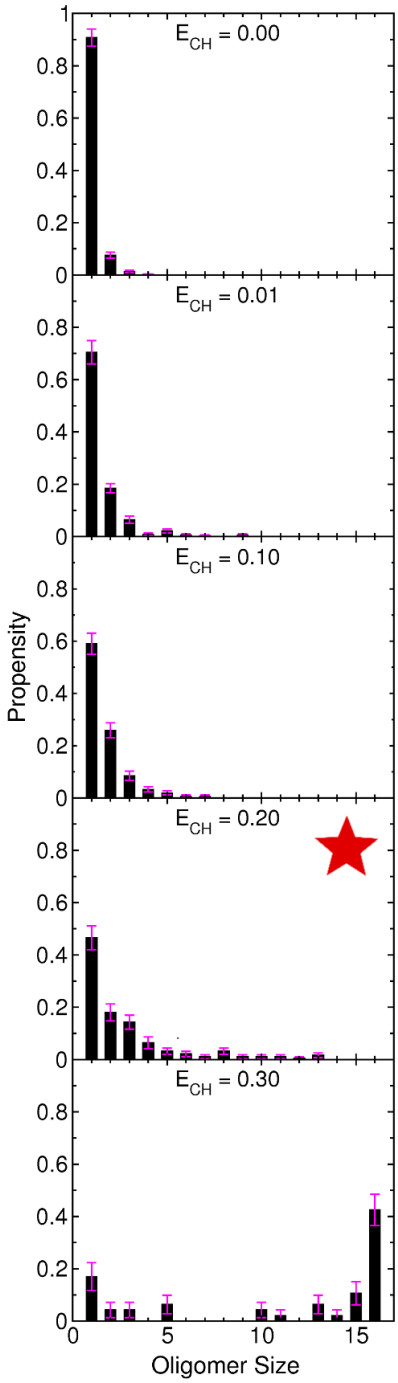

Figure 4.2: Effect of E_{CH} value on the oligomer size distribution. Oligomer size propensities for simulated E_{CH} values ranging from 0.00 to 0.30 as observed at the final time step of DMD simulations. Conformations from simulations with $E_{CH} = 0.20$ (red star) were chosen for further analysis.

We compared our oligomer distributions for each E_{CH} value to oligomer size populations that have been experimentally derived primarily by PICUP and SDS-PAGE[199–204;207]. This experimental method is better suited for detection of LMW rather than higher order oligomers, however, it is currently the only technique that offers insight into the relative propensity of *in vitro* and *in vivo* oligomer populations. Multiple studies by Ono et. al. reported a mixture of wild-type (WT) αS monomers, dimers, trimers, and tetramers when using an αS concentration of 70 μM but did not report on their relative propensities[199–201]. Ghosh et. al. reported monomeric to pentameric αS species at a concentration of 300 μM, whereby the SDS-PAGE abundance pentamer band appears more dense than the tetramer band, indicating a larger abundance of the former[203]. Acharya et. al. characterized the oligomer size distribution at both 3 μM and 30 μM concentrations[204]. At 3 μM, their αS oligomer size distribution was dominated by monomers and dimers, although trimers, tetramers, and pentamers were noted as well. At 30 μM, a significant reduction in monomers with a corresponding increase in dimers although pentamers was observed relative to the 3 μM concentration[204]. Given that our *in silico* concentration of αS was $\sim$3 mM, we favored results derived at a experimental concentration for comparison to our data. Acharya et. al. also quantified αS monomers and oligomer abundances, and their normalized oligomer size distribution at 30 μM is comparable to the DMD4B-HYDRA-derived oligomer size distribution for $E_{CH} = 0.20$. It is important to note, however, that our DMD-derived distribution contained larger oligomer sizes which were not reported by Acharya et. al. There is evidence of αS oligomers larger than pentamers being present *in vitro*; Borsarelli et. al. observed a monotonically decreasing propensity of αS monomers through tetramers for PICUP irradiation times up to 200 seconds, as well as the presence of 'high molecular weight oligomers, at a concentration of 100 μM[202]. Although they were not able to specifically characterize these larger oligomer sizes due to the resolution of their SDS-PAGE results, they determined the 'mean species' to be that of a hexamer[202]. Dettmer et. al. also reported assemblies which were $\sim$ 80kDa and $\sim$100 kDa, weights consistent with αS hexamers and septamers, although they were not able to definitively determine if these assemblies were not merely smaller oligomers in a more extended conformation[207].

When taking into account that the concentration of αS in our simulations is $\sim$ 3mM, which is one to two orders of magnitude higher than any of the experimental SDS-PAGE results described above, we expect to observe the presence of larger oligomer orders relative to experimental data. Collectively, the above considerations led us to select DMD4B-HYDRA trajectories obtained at E_{CH} = 0.20 for further analysis. Monomers through tetramers were reported in each of the SDS-PAGE results described above, with pentamers also directly reported by Ghosh et. al. and Acharya et. al. These results motivated us to structurally analyze monomers, dimers, trimers, tetramers, and pentamers in the present study. We also explored oligomer sizes beyond pentamers, as the presence of such assemblies was reported by Borsarelli et. al.[202] In the DMD4B-HYDRA-derived oligomer size distribution for $E_{CH} = 0.20$, octamers were the only relatively abundant oligomer size larger than pentamers. We thus included octamers in our structural characterization of αS monomers and oligomers.

4.2.2 Convergence of the Oligomer Size Distribution

When studying IDPs, it is important to account for the amount of time required for the population of oligomers to reach a metastable state. We explored the time evolution of the oligomer size distribution (Figure S1 in Appendix B) in order to restrict structural analysis to time frames that displayed relatively stable αS oligomer populations. As expected, the number of monomers diminished with time. Monomers, dimer, tetramer, pentamer, hexamer, and septamer propensities did not fluctuate beyond the calculated error bars from 40M time units onward. Trimers and octamers, however, did not reach stable propensities until 60M time units. From 60M to 80 M time units, the propensity for monomers and oligomer sizes up to and including octamers did not change significantly. We thus selected simulation frames between 60M and 80 M time units from simulations with $E_{CH} = 0.20$ for subsequent analyses. The exceptions were the PMFs and reaction coordinate distributions, which utilized all conformations from 0-80 M time units. The number of monomer and oligomer conformations present in simulations frames from 60M and 80M time units and 0-80 M time units is presented in Table S1 in *Appendix B*.

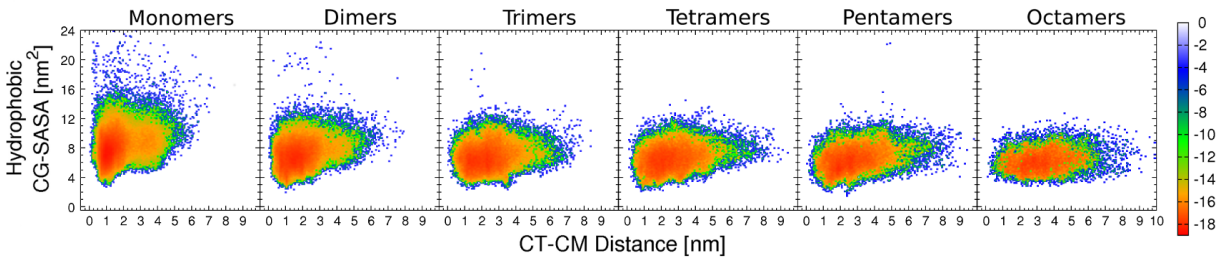

Figure 4.3: Free energy landscapes. PMF plots of αS conformations projected onto two reaction coordinates, the CT-CM distance and hydrophobic CG-SASA. Values were calculated for each individual peptide from each oligomeric conformation. The color scale on the right is given in units of k_BT

Free Energy Landscapes and Reaction Coordinate Distributions

Potential of mean force (PMF) plots (Figure 4.3) illustrate the free energy landscape of a peptide, allowing insight into the most often assumed conformations. We chose to include all conformers from the enire 80 M time unit range of simulation to allow for the largest possible exploration of the phase space, as defined by the two reaction coordinates: the distance from the C-terminal of each peptide (A140) to the center of mass (CM) of the monomer or oligomer (CT-CM distance) and the sum of the coarse-grained (CG) solvent accessible surface area (SASA) of all hydrophobic residues, as defined by the Kyte-Doolittle scale[3] (Hydrophobic CG-SASA). The free energy landscapes in Figure 4.3 shifted downward with increasing oligomer order, as shielding of hydrophobic residues from the solvent increases with oligomer order. We chose to closely examine the C-Terminal position relative to the CM because the CTR is most hydrophilic region of the αS sequence and is thus energetically most likely to be exposed to the solvent. In Figure 4.3 the free energy landscapes expand to larger CT-CM distances with increasing oligomer order. While this observation can be in part attributed to an increase in radius with increasingly larger oligomers, it informed us that A140 does not remain close to the core of the oligomer as oligomer size increases.

One-dimensional probability distributions of the CT-CM distance and hydrophobic CG-SASA reaction coordinates can be seen in Figure 4.4, with average values of the distributions displayed in Table 4.1. The reduction in the hydrophobic CG-SASA is largest when going from monomers

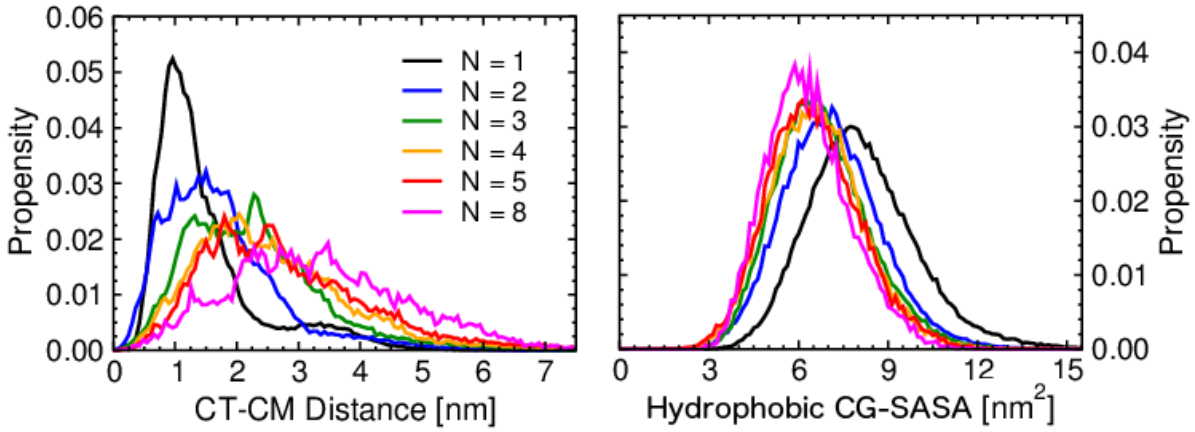

Figure 4.4: Effect of assembly state on distributions of CT-CM distance and hydrophobic CG-SASA. Mann-Whitney U tests comparing all assembly size distributions resulted in $p < 10^{-3}$ for all.

to dimers, after which the reduction is not as large with increasing oligomer order. The observed hydrophobic collapse is a driving force for protein aggregation, and the dramatic reduction in hydrophobic CG-SASA with dimerization reflects this idea. The fleeting reduction of hydrophobic CG-SASA with increasing oligomer order, however, implies the importance of other driving forces in the formation of larger oligomers. The CT-CM distance distributions shifted to larger values with increasing oligomer order, as did the N-terminal to CM (NT-CM) distance (Figure 4.5). The average CT-CM distance was significantly larger than the average NT-CM distance for all assembly sizes, reflecting the influence of the increased hydrophobicity of the NTR relative to the CTR (Figure 4.5). The CT-CM distance distributions also attain larger values than the NT-CM distance distributions for all assembly sizes.

The final reaction coordinate we explored was the N-terminal to C-terminal (NT-CT) distance (Figure 4.5). The average NT-CT value increased with oligomer size but was not significant when going from trimers to tetramers or from tetramers to pentamers (Table 4.1). NT-CT distributions revealed an interesting characteristic of αS peptides in our simulations; two distinct peaks in the distribution, which correspond to conformations with the lowest free energy, are present for every assembly size, centered around $\sim 4-6$ Å (cluster 1) and $\sim 12-17$ Å (cluster 2). Hydrophobic CG-SASA values for these both of these free energy minima span $\sim 50-95$ Å^2. As these free

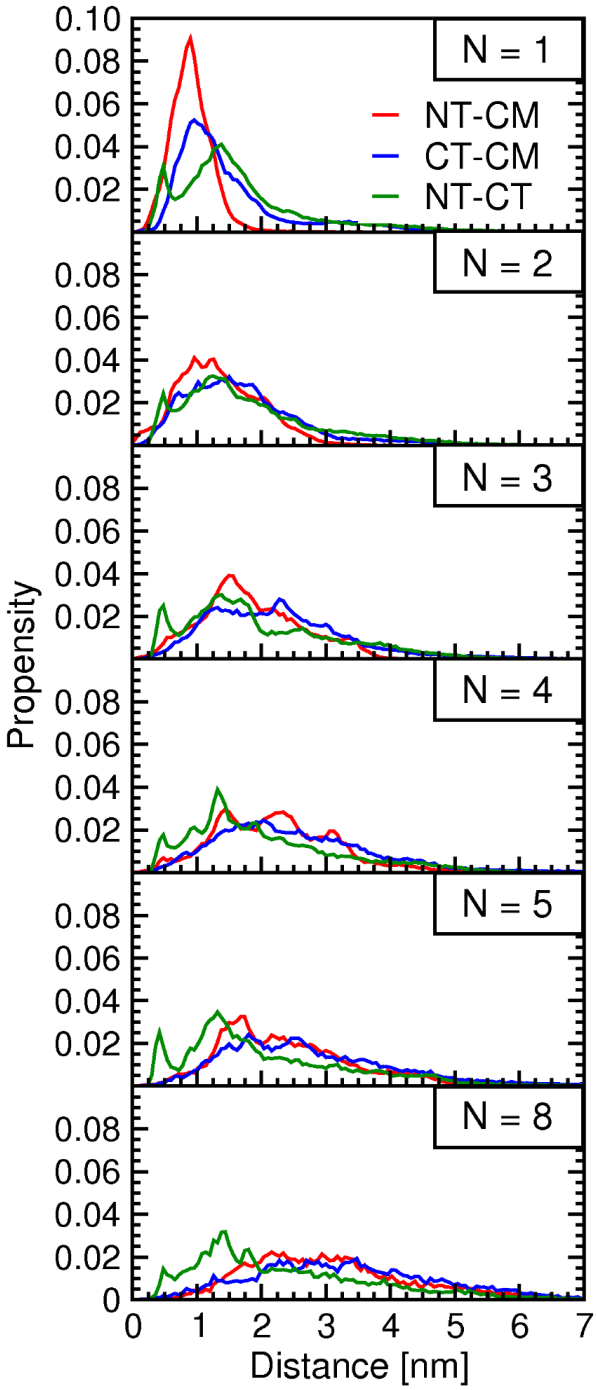

Figure 4.5: Comparison of NT-CM, CT-CM, and NT-CT distance distributions for each assembly state. Mann-Whitney U tests comparing distributions resulted in $p < 10^{-3}$ for all assembly sizes.

energy landscapes were generated by creating a data point for each peptide in a given oligomer (expanded upon in the *Methods* section), both monomers and the individual peptides which make up the oligomers in the present study are characterized by this feature.

Table 4.1: Average values of reaction coordinate probability distributions. Mean values correspond to the distributions displayed in Figure 4.4. Errors in the mean values correspond to SEM values.

N	NT-CM [nm]	CT-CM [nm]	NT-CT [nm]	CG-SASA [nm^2]
1	0.91 ± 0.01	1.49 ± 0.02	1.70 ± 0.02	8.15 ± 0.03
2	1.36 ± 0.02	1.73 ± 0.02	1.84 ± 0.02	7.12 ± 0.04
3	1.85 ± 0.02	2.27 ± 0.03	2.02 ± 0.03	6.78 ± 0.04
4	2.21 ± 0.03	2.54 ± 0.03	2.06 ± 0.03	6.61 ± 0.03
5	2.37 ± 0.03	2.77 ± 0.04	2.11 ± 0.04	6.52 ± 0.05
8	3.08 ± 0.06	3.34 ± 0.07	2.36 ± 0.06	6.33 ± 0.07

In order to gauge the extent to which electrostatic interactions influenced this feature, we created the same NT-CT probability distributions for the assembly sizes present in simulations with E_{CH} set to 0.00 (Figure S2 in *Appendix B*), which included monomers, dimers, and trimers. We observed that the unique peaks in the probability distributions from simulations with $E_{CH} = 0.20$ were completely absent when electrostatic interactions were turned off, replace by a gaussian-shaped distributions centered around $\sim 35 - 45$ Å for all assembly sizes. We conclude therefore that the electrostatic interactions are essential for the existence of this feature.

4.2.3 Unique Monomer Conformations Inhabit Distinct Free Energy Minima

We separated and explored two populations of monomers within the two distinct peaks in the NT-CT distance probability distribution in order to gain insight into other potential defining characteristics (Figure 4.6). We selected the ∼1,000 conformations from cluster 1 (black ellipse) and ∼2,500 conformations from cluster 2 (white ellipse) with the lowest free energies. Cluster 2 covered a larger portion of the overall landscape and thus contains more conformations. Three characteristic conformations

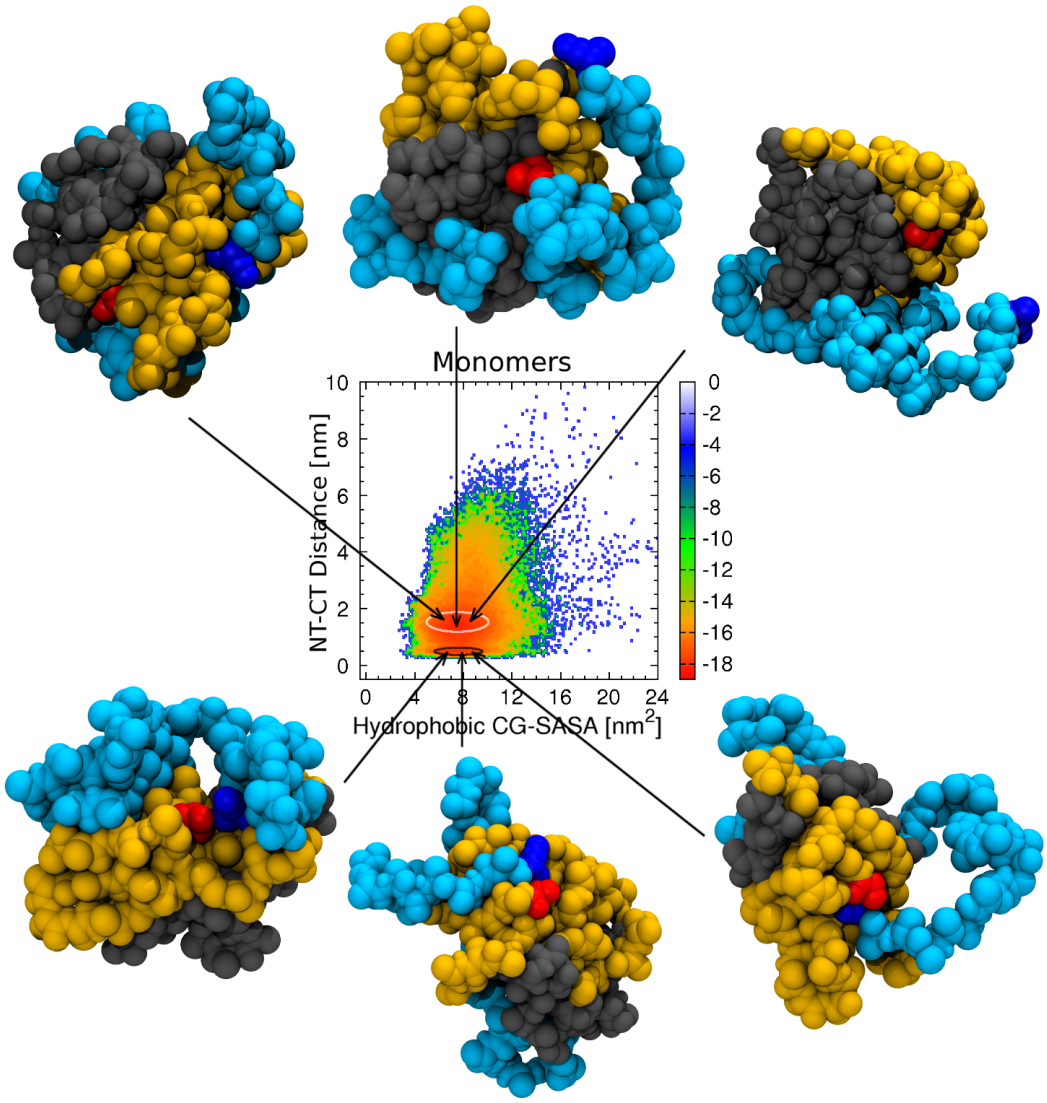

Figure 4.6: Representative monomer conformations. Monomeric αS conformations were projected onto two reaction coordinates, NT-CT distance and hydrophobic CG-SASA. The displayed monomer conformations were drawn from cluster 1 (black ellipse) and cluster 2 (white ellipse). The C-terminus is represented by blue spheres and the N-terminus by red spheres. The NTR (yellow spheres), NAC (grey spheres), and CTR (cyan spheres) are also displayed.

from each cluster are displayed in Figure 4.6. Conformations from both clusters had a similarly globular morphology with the NAC and NTR regions forming the core, and the CTR wrapping around the surface of the core. In all cluster 1 conformations, M1 and A140 are in contact with each other, and these two residues were found on or very near the surface of the peptide. This is distinct from cluster 2 conformations, in which M1 and A140 were not in contact and M1 was frequently buried within the core of the peptide, and thus less accessible to A140. Cluster 2 conformations occasionally featured a completely solvent exposed C-terminus (displayed in Figure 4.6), but much more frequently A140 was found on the surface of the peptide. In cluster 1 conformations, A140 was often buried in the NAC or NTR. We plotted the solvent exposure and arrangement of residues relative to the CM of monomers from clusters 1 and 2 in Figure 4.7. M1-F4 are on average more exposed to the solvent and found farther from the CM on average in cluster 1 conformations than in cluster 2 conformations (Figure 4.7), while A140 is less solvent exposed and closer to the CM in cluster 1 than in cluster 2 conformations. This feature is important because it affects the CTR conformation. The CTR is the most hydrophilic region of the peptide sequence, and at what point in the sequence A140 'anchors' itself to the surface of the globular conformation affects the rest of the CTR. In cluster 1 conformations, in which M1 and A140 are in contact, the N103-Q109 and D120-S129 regions are on average farther from the CM (Figure 4.7, bottom panel) than in conformations from cluster 2. The N103-Q109 and D120-S129 regions are frequently a part the loop that forms a porelike structure in conformations from both cluster 1 and cluster 2 in Figure 4.6. Consequently, most of these two regions are more exposed to the solvent in cluster 1 than in cluster 2 conformations (Figure 4.7, top panel). The Y133-E139 region was also frequently a part of porelike monomer structure, and this region was found farther from the CM and were more solvent exposed in cluster 2 than in cluster 1 conformations. On average, the diameter of porelike structure in cluster 1 conformations was larger than those in cluster 2 conformations. Though pores occurred with relatively equal propensity in both cluster 1 and cluster 2, conformations from cluster 1 were more likely to have a pore diameter that was at least $\sim$3 Å in diameter.

Secondary structure was also substantially different between monomer conformations from these

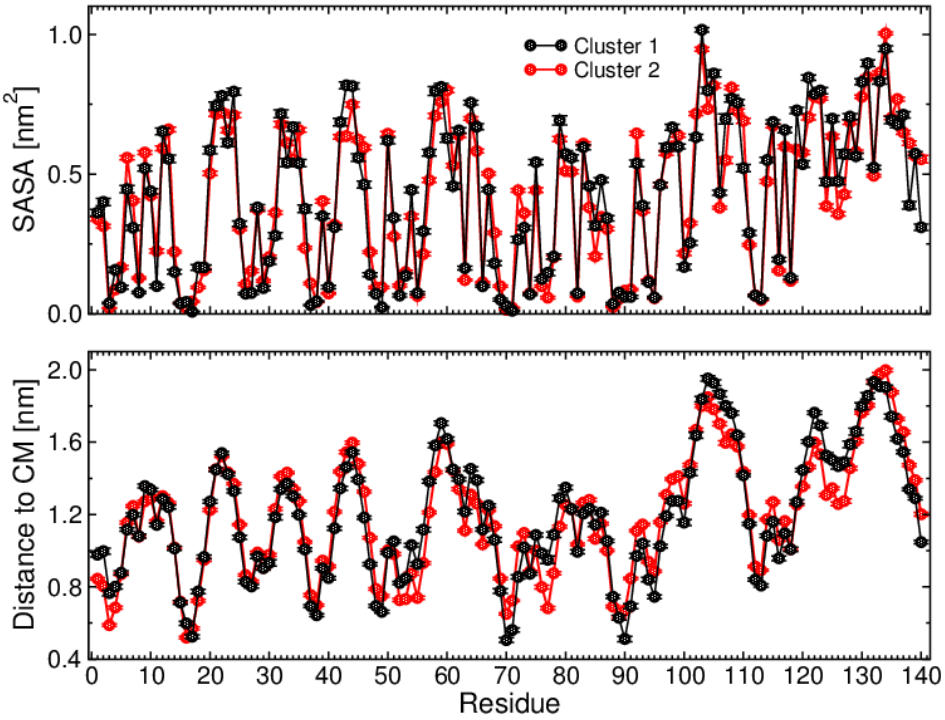

Figure 4.7: Average per-residue solvent accessible surface area and distance to CM for monomer conformations from clusters 1 and 2. Error bars correspond to SEM values.

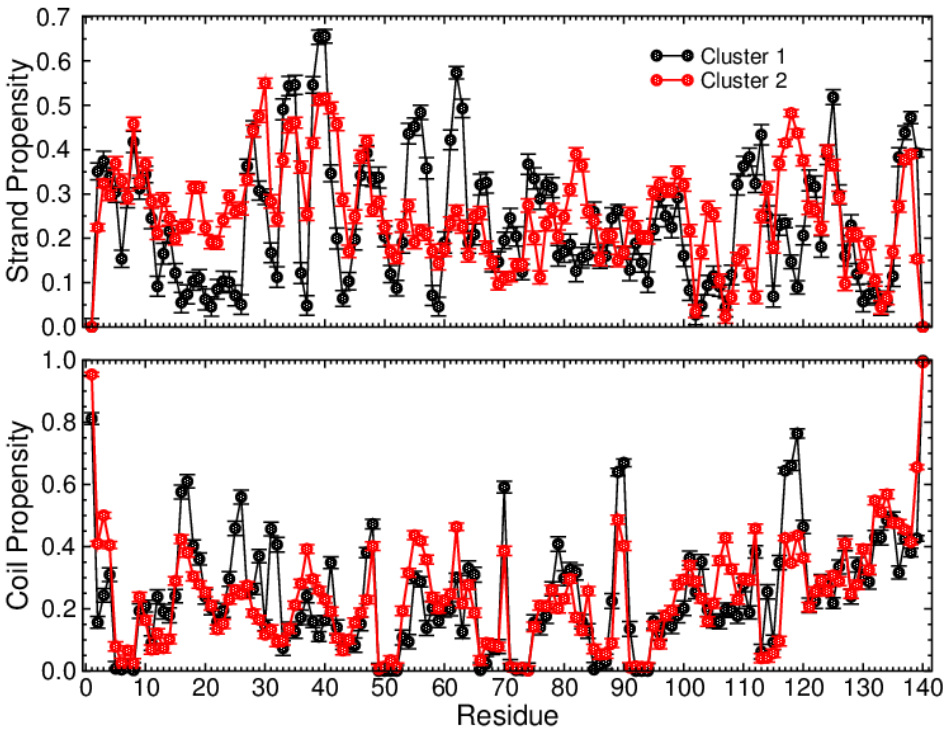

Figure 4.8: Average per-residue β-strand and statistical coil propensities for monomer conformations from clusters 1 and 2. Secondary structure was obtained via the STRIDE program in VMD. Error bars correspond to SEM values.

two clusters on a per-residue basis. This is most notable in the V15-V26 region, which had $\sim 15\%$ more β-strand in cluster 2 conformations compared to those from cluster 1 (Figure 4.8). The E110-P120 region had significant (up to $\sim 30\%$) differences in β-strand content between conformations from the two clusters, whilst cluster 1 conformations were significantly less structured overall across these residues. Cluster 1 conformations had significantly more turn content in the N103-Q109 region (Figure S3 in *Appendix B*), which was recognized above to frequently contribute to pore-like conformations. The Y136-E139 region had notably more β-strand content in cluster 1 than in cluster 2 conformations, reflecting the importance of which contact A140 makes with the rest of the peptide. Although differences in the average β-strand propensity exist across smaller regions of the primary structure, the average β-strand, turn, bridge, and statistical coil propensities were comparable for clusters 1 and 2 conformations (data not shown).

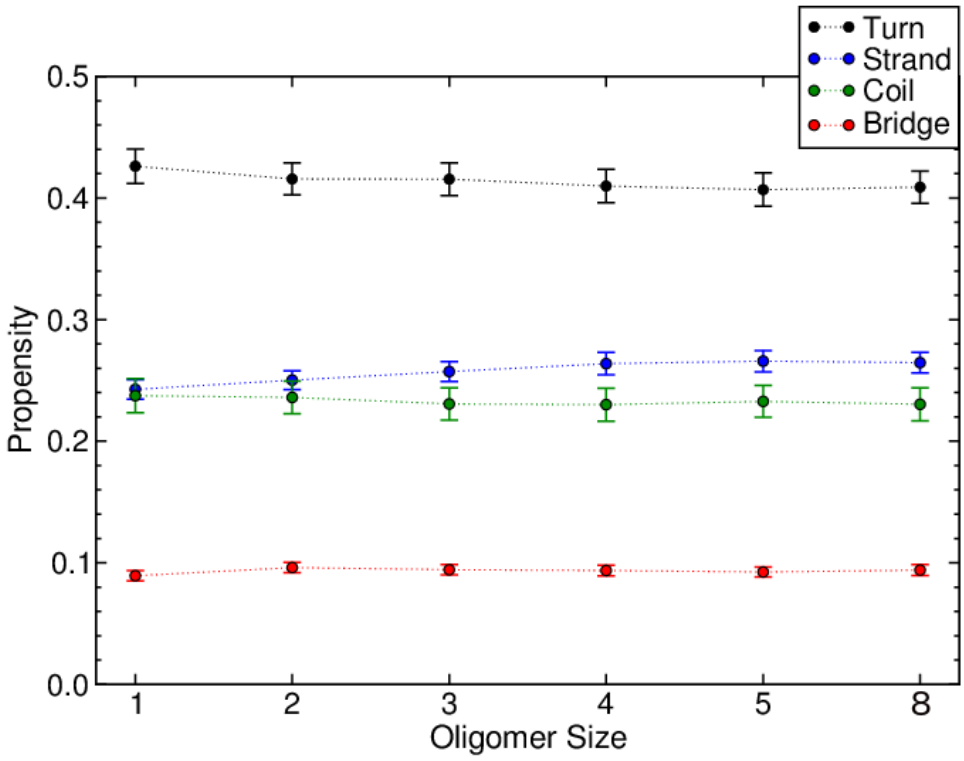

Figure 4.9: Average secondary structure of αS conformations. Values were averages over each peptide. Error bars correspond to SEM values.

4.2.4 Turn Content is the Predominant Secondary Structure in αS

We examined the secondary structure of monomer and oligomer conformations using STRIDE[157] implemented within the Visual Molecular Dynamics (VMD) software[169]. Calculation of the average secondary structure per peptide (Figure 4.9) revealed that turn content was the predominant secondary structure across all assembly sizes ($\sim$ 40-44%), followed by β-strand ($\sim$ 23-28%) and statistical coil ($\sim$ 22-25%), bridge ($\sim$ 8-10%), and α-helix ($\sim$ 1-2%). Average β-strand and coil was comparable for monomer and dimer conformations, whereas β-strand becoming was prominent than coil in assembly sizes larger than dimers. Average secondary structures did not significantly change with assembly state up to trimers, whereas tetramer, pentamer, and octamer conformations had more average β-strand content per peptide than monomer conformations.

The average secondary structure per residue revealed region-specific secondary structure features

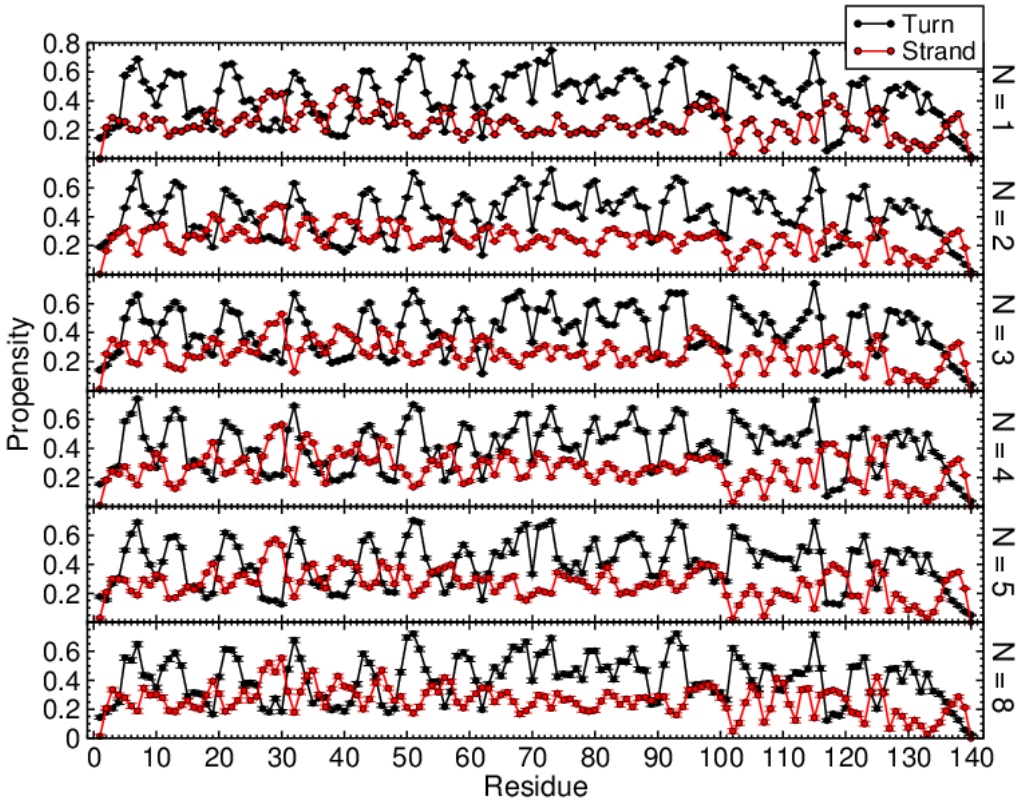

Figure 4.10: Average per-residue turn and β-strand content of αS conformations Error bars correspond to SEM values.

(Figure 4.10). The NAC region had the highest turn content of the three regions. β-strand content exceeded turn content across 3 or more amino acids in all assembly states in the A27-A30, L38-G41, E46-V48, P117-P120, and E137-E139 regions. Turn content dominated all other peptide regions. Amino acids E28-A30 had the highest β-strand propensity per-residue ($\sim$ 40-60%) in the entire sequence in all assembly states, as well as in the Y39-G41 region in monomers. Most of the residues in the two regions associated with the porelike structures observed of monomers, N103-Q109 and D120-S129 (Figure 4.6), had significantly more turn content than β-strand content in all observed assembly states. Residues E137-E139, however, had notably more β-strand than turn content. The CTR was the least structured region of the peptide sequence for monomers and oligomers, containing the least β-strand content (Figure 4.10) and the most coil content of the three major regions of the peptide sequence (Figure S4 in *Appendix B*).

4.2.5 Tertiary and Quaternary Structure

To understand the tertiary and quaternary structure of αS, intramolecular (Figure 4.11A) and intermolecular (Figure 4.11B) contact maps between pairs of C_β atoms were created for all assembly sizes under present study as defined in *Methods*. Contact maps between pairs of C_α atoms were calculated as well (see Figure S5 in *Appendix B*). Both intra- and intermolecular contact maps are split up into six distinct sectors to elucidate contacts between and among the three defined regions of the peptide sequence; NTR-NTR contacts (sector 1), NTR-NAC contacts (sector 2), NTR-CTR contacts (sector 3), NAC-NAC contacts (sector 4), NAC-CTR contacts (sector 5), and CTR-CTR contacts (sector 6). Contacts in a given sector of the intramolecular contact maps occur between residues from the same peptide. Contacts in a given sector of the intermolecular contact maps occur between residues belonging to the corresponding region or regions of different peptides.

Intramolecular contacts which occur close to the diagonal are the strongest and contribute to the stability of local structure. Off-diagonal contacts occur between peptide regions that are not directly adjacent; the farther from the diagonal, the farther apart the residues in contact are in the primary structure. C_β-C_β contact maps, both intramolecular and intermolecular (Figure 4.11), contain a larger number of high propensity contacts than C_α-C_α contact maps (Figure S5 in *Appendix B*). The contacts described below are features of C_β-C_β contact maps unless otherwise noted.

Tertiary Structure

Intramolecular contacts did not vary greatly between assembly sizes, particularly after dimerization (Figure 4.11A). The strongest intramolecular NTR-NTR contacts occur between the V3-M5 and G15-E18 regions in both C_β-C_β (Figure 4.11A, sector 1) and C_α-C_α (Figure S5A in *Appendix B*, sector 1) contact maps, and these contacts are persistent across all assembly sizes. The G15-E18 region also made strong contacts with NTR residues K21-T22, V37-L38, V48-V49, K58, and K60 in all assembly sizes. These contacts were notably weaker in C_α-C_α contact maps (Figure S5A in *Appendix B*, sector 1), suggesting their formation was largely driven by effective electrostatic interactions between amino acid side-chains. Sector 1 of the C_α-C_α contact maps contains distinct complexes of anti-parallel tertiary contacts which are present in all assembly sizes. The first β-

hairpin occurred between residues A27-A30 and K34-G36 with a turn across residues G31-T33, and a second β-hairpin between residues L38-G41 and K45-V48 with a turn across S42-T44. The location of these anti-parallel features align well with regions of increased β-strand propensities (Figures 4.9 and 4.10).

The NTR-NAC intramolecular contacts (Figure 4.11A, sector 2) are strongest in monomers, particularly between NTR residues V48-V49, V51-A52, and T54, and NAC residues V70-V71, V74, and A76-V77. Most of these contacts diminished in strength upon dimerization and remained weaker relative to monomers in all oligomeric states. Any contacts involving NAC residues from V66 to V82 are important to note, as this hydrophobic sequence has been shown previously to play a crucial role in the formation and structure of αS fibrils[93;208]. The NTR-CTR contacts are overall weaker than the NTR-NTR and NTR-NAC contacts (Figure 4.11A, sector 3). The strongest NTR-CTR contacts occurred between the G15-E18 and I112-L113 regions in monomers, and were weaker in oligomers. A140 has the highest contact propensity with M5, L8, K11, and V15 in monomers, and these contacts are again weaker in oligomers. These were also the frequent contacts formed by A140 in monomers from cluster 2. Prominent NTR-NAC and NTR-CTR contacts were weaker overall in C_α-C_α contact maps (Figure S5A in *Appendix B*, sectors 2 and 3). relative to C_β-C_β, indicating that hydrophobicity plays a key role in their formation.

The NAC-NAC intramolecular contacts are the strongest contacts observed, both diagonal and off-diagonal, in both C_β-C_β and C_α-C_α contact maps (Figure 4.11A, and Figure S5A in *Appendix B*, sector 4). Residues V66 and A69-V71 formed strong contacts with the V77-Q79 region across all assembly states. Residues V66, A69-V71, V74, and A76-A78 also form prominent off-diagonal contacts with the I88-A91 region which remain strong across all assembly states. These NAC-NAC contacts in sector 4 reflect a strong propensity for the NAC to collapse onto itself, which stabilized the core of the globular morphology in monomer conformations (Figure 4.6). This also resulted in the the high turn content in the NAC in Figure 4.10.

The NAC-CTR intramolecular contacts are the most prevalent in monomers, and are stronger in the C_β-C_β contact maps (Figure 4.11A, and Figure S5A in *Appendix B*, sector 5). Residues

I112-L113 formed the strongest C_β-C_β contacts in sector 5, with the NAC residues V70, V77, I88-A90, and V95 (Figure 4.11A, sector 5). These contacts are present in oligomers but are not as strong, and were particularly inhibited in octamers. I112-L113 represented the most frequent point of intramolecular contact between the CTR and the NTR/NAC.

Quaternary Structure

Quaternary contacts, which occur between residues of different peptides, contribute to aggregation and oligomer stability. Sector 1 of the intermolecular contact maps displays contacts between NTR pairs (Figure 4.11B, and Figure S5B in *Appendix B*). Strong contacts occurred within the first 20 residues of the sequence; M1 and the V3-M5 region made strong contacts with V3-M5, L8, V15-A19, and V48-V49. Residues V15-A19 also made strong contacts with V15-A19, particularly in trimers and pentamers, as well as with residues V48-V49, V52-A53, and V55. These intermolecular contacts spanning sector 1 were comparable across oligomer sizes but notably diminished in octamers. V49-V49, V49-V52, and V52-V52 also formed strong intermolecular NTR-NTR contacts in dimers, but are weaker in larger oligomers. Nearly all of these residues are hydrophobic, stressing the importance of hydrophobicity in driving oligomer formation. NTR-NTR contacts were mostly comparable or stronger in C_β-C_β than in C_α-C_α contact maps (Figure 4.11B, and Figure S5B in *Appendix B*, sector 1).

The NTR-NAC intermolecular contacts (sector 2) are weaker and less numerous than those in sector 1 (Figure 4.11B, and Figure S5B in *Appendix B*). NTC-CTR contacts are some of the strongest and most numerous intermolecular contacts across all oligomer orders (Figure 4.11B and Figure S5B, sector 3). Residues I112-L113 were frequently in contact with the V15-A18 region and are the strongest in trimers and octamers. The D135-E139 region made contacts with the K32-K34 region in all assembly sizes and with the S42-K45 region in trimers and larger oligomers, whereby the strength of these latter contacts increased with oligomer order. In addition, the D135-E139 region formed strong contacts with the K10-K12 and E20-K23 region in pentamers and octamers. A140 formed the most numerous strong intermolecular contacts of any residue in the sequence. A140 contacted the F4-K6, A11-K12, and V15-A19 regions in all assembly sizes, although they

were strongest in pentamers and octamers. In trimers and larger oligomers, A140 also made strong contacts with V37-V40, K43-T44, and V55-A56. These contacts are again strongest in pentamers and octamers. Sector 3 contacts were largely comparable in strength or stronger in C_β-C_β than in C_α-C_α contact maps (Figure 4.11B, and Figure S5B in *Appendix B*, sector 3).

The NAC-NAC intermolecular contacts (Figure 4.11B, sector 4) are of particular importance, as the NAC region is knon to be involved in αS fibril formation[92–94;208]. Contacts in this sector overall increase in strength with oligomer size. The NAC-NAC contacts were weak in dimers and trimers. In tetramers, V70-A89 was the strongest occurring contact, followed by those between the S87-A89 and A88-A89 regions, which were weaker in pentamers but stronger in octamers. The A69-V70 region formed the most numerous intermolecular contacts in sector 4 of pentamers, the strongest of which were involve the A69-V70 region and F94 (Figure 4.11B), also present in octamers. Octamers had the strongest NAC-NAC intermolecular contacts. In addition, there is a group of strong contacts between A69-A78 regions.

The CTR does not form strong intermolecular contacts with the NAC, with the exception of A140 with A89 and V95 in tetramers, with A78 in pentamers, and with A85 in octamers. The contact A89-G111 was also prominent in octamers, but only in C_β-C_β contact maps (Figure 4.11B, and Figure S5B in *Appendix B*, sector 5). The CTR-CTR intermolecular contacts are the weakest and least numerous between any regions of the primary structure. Similar to the rest of the intermolecular contacts, C_β-C_β contacts were overall stronger than or comparable in strength to C_α-C_α contacts in sectors 4, 5, and 6 (Figure 4.11B, and Figure S5B in *Appendix B*).

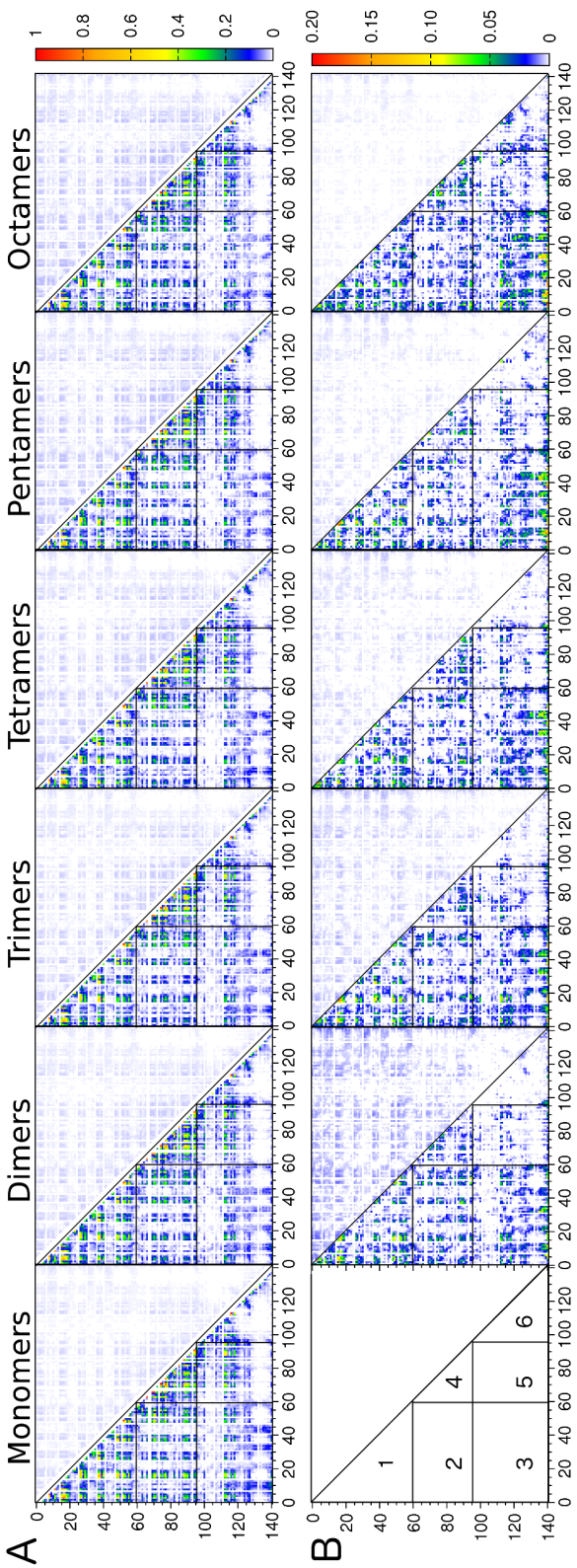

Figure 4.11: C_β-C_β contact maps. (A) Intramolecular and (B) intermolecular contacts are plotted below the diagonal with corresponding SEM values above the diagonal. The color scale on the right shows the contact strength as defined in the *Methods* section. Contact maps are divided into 6 sectors (bottom left).

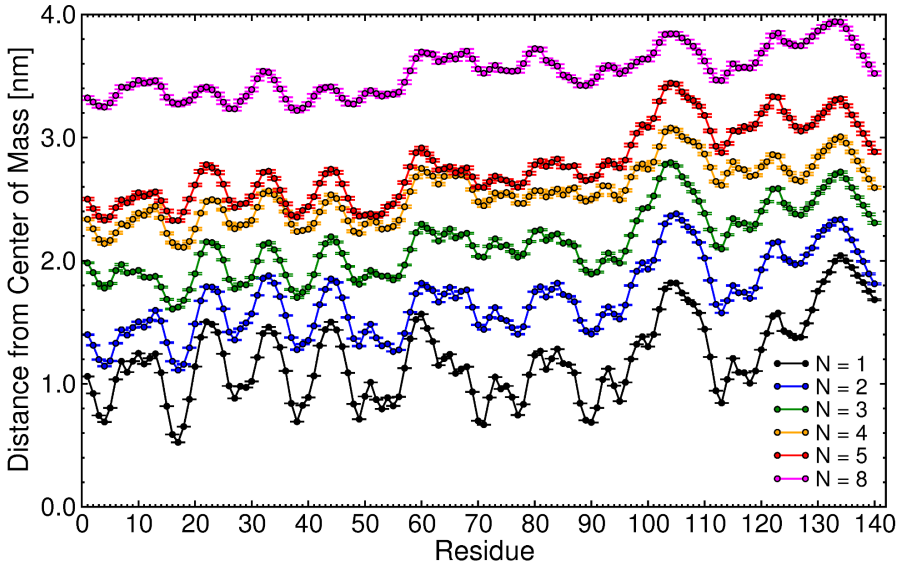

Figure 4.12: Average distance from the CM per amino acid: assembly state effect. Error bars correspond to SEM values.

4.2.6 Residue-Specific Arrangement and Solvent Accessibility

The arrangement of amino acids relative to the CM of the monomer or oligomer is presented in Figure 4.12. The distance from the CM across the sequence increases with oligomer order, however this increase was notably lessened between tetramers and pentamers for much of the NTR and NAC. The CTR was the farthest from the CM for all assembly sizes. We also observed that for all assembly sizes, the farthest residues from the CM in the CTR, and overall, coincided with regions of the sequence which were frequently observed to be a part of porelike conformations in monomers (N103-Q109, D120-S129, and Y133-E139, Figure 4.6). The difference in the average distance from CM between the CTR and the other regions of the peptide sequence increased with oligomer order up to and including pentamers. In pentamer conformations, nearly every CTR residue was more distant from the CM than every NTR and NAC residue. In monomers, the NAC was on average closer to the CM than the NTR; these two regions form the core of the monomer morphology as displayed in Figure 4.6. Upon dimerization, the NAC moved farther from the CM than the NTR, an effect which increases with oligomer order. In octamer conformations, the NAC and CTR regions were at comparable distances from the CM.

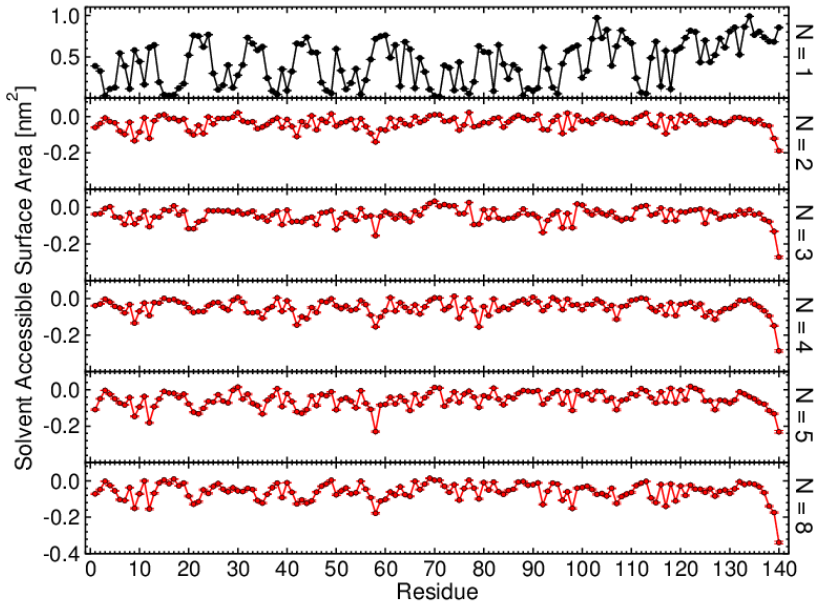

Figure 4.13: Average SASA per amino acid relative to monomeric αS. SASA is plotted for monomers and difference in SASA relative to monomers is plotted for all oligomer sizes. Error bars correspond to SEM values.

We also explored the average SASA per residue, and found the CTR to be the most solvent exposed region of the peptide in all assembly states (Figure S6 in *Appendix B*). Residues I112-L113 had the lowest SASA in the CTR, which is to be expected following the observation that these residues frequently made strong contacts with the other hydrophobic peptide regions (Figure 4.11, and Figure S5 in *Appendix B*). The regions associated with porelike morphologies in monomers, N103-Q109, D120-S129, and Y133-E139 (Figure 4.6), had relatively high SASA values. This feature was persistent across assembly states. We calculated the difference in SASA between monomers and oligomers to observe the impact of oligomerzation on SASA values across the entire sequence. In all assembly states, A140 was associated with the largest SASA reduction relative to monomers, followed S9, K58, and E139. In pentamers and octamers, K12 also had a significant decrease in SASA. The solvent exposure of P138 also decreased with increasing oligomer order. The amino acids which displayed increased solvent exposure were largely hydrophobic residues, however the increase was very small $<.05\ nm^2$).

4.2.7 Morphological Heterogeneity Increases with Oligomer Orders

In order to explore morphologies of αS monomers and oligomers in our simulations, we calculated the principal moments of interta ($I_1 \geq I_2 \geq I_3$) and created a phase space with two ratios, I_2/I_1 and I_3/I_1 (Figure 4.14). All data points lie below the diagonal line, as by definition $I_2 \geq I_3$. Conformations which lie near the diagonal in the upper-right portion of the plot tend to be more spherical in shape ($I_1 \approx I_2 \approx I_3$). Conformations near the bottom-right of the plot, as I_3/I_1 approaches zero and $I_1 \approx I_2$, tend to be elongated and rod-like. Curvilinear morphologies, such as annular, bent ("L" and "V"-shaped), and "Y"-shaped conformations, are shifted to the left side of the plot, where $I_2 \approx I_3$.

Examples of monomer and oligomer morphologies which were present at the end of 80 M time units are presented on either side of Figure 4.14 for each assembly state, and represent some of the morphologies present in each assembly size. Conformations on the left side of Figure 4.14 feature a globular monomer and curvilinear oligomer morphologies. On the right side, a more elongated but still globular monomer is presented along with rod-like and elongated oligomer morphologies. Examples of oligomer morphologies are also displayed in Figure 4.15. Monomers were largely globular and the most spherically shaped assembly size. The relative positions and conformation within the CTR can influence the morphology of monomers, as the CTR can extend relatively far from the CM as seen in Figure 4.14 (right side). Upon dimerization, morphologies shifted towards more extended and explicitly less spherical conformations, a trend which continues up to an including tetramers (Figure 4.14 right side, Figure 4.15B). Both dimers and trimers also adopted bent, curvilinear morphologies (Figure 4.14 left side), but these were not as common as extended morphologies. The variety of morphologies expanded in oligomers larger than trimers. Bent (Figure 4.14 left side), "Y"-shaped (Figure 4.15F and G) and ring-like, annular morphologies (Figure 4.15E) became prominent in tetramers, although elongated conformations were still the most abundant (Figure 4.14 right side). In pentamers, extended rod-like conformations (Figure 4.14 right side), and curvilinear, bent (Figure 4.14 left side) and "Y"-shaped morphologies (Figure 4.15I,J) were occur with more equal probabilities than in tetramers. Pentamers also displayed annular morphologies, similar to tetramers (Figure 4.15H). This trend continues in octamers, as bent and curvilinear (Figure 4.14 left side, Fig-

ure 4.15L), but not "Y"-shaped, morphologies became more typical than the extended conformations. Elongated octamers appeared as loosely connected smaller oligomers (Figure 4.14 right side), in contrast to the more cylindrical rod-like conformations, which were observed in smaller oligomers. The annular morphology observed in tetramers and pentamers was also present in octamers. Additional planer, amorphous octamer conformations that were roughly disk-like (Figure 4.15K,M) were also observed. These results showcase the polymorphic nature of αS oligomers, which is prominent in oligomers larger than trimers.

Regardless of morphology, exploration of αS conformations confirms the importance of the CTR in LMW oligomer formation. We observed the CTR to make numerous strong intermolecular contacts in oligomers of all sizes (Figure 4.11B, and Figure S5B in *Appendix B*). Throughout the conformations presented in Figure 4.15, the CTR can be seen forming contacts with residues from other peptides. Figure 4.15M provides the clearest example, as this octamer is comprised of a monomer, trimer, and tetramer connected through their several intermolecular CTR contacts.

4.2.8 Porelike αS Morphologies

The leading hypothesis concerning the root cause of αS oligomer-induced toxicity, the ion channel hypothesis, states that αS assemblies are capable of forming annular pores in a cellular membrane which act as ion channels, resulting in an abnormal flux of ions and destabilizing the homeostasis of the cell[83;100–102;104;107;111–113;209]. We observed porelike structures in monomers which were always composed primarily of residues in the CTR, the most hydrophilic region of the peptide. To determine if this characteristic was present in oligomers, we examined all conformations present in the final time unit of simulation time for the presence of pores. We defined a 'pore' to be a closed loop structure with a diameter of at least $\sim$3 Å. We observed the presence of porelike structures in conformations from all oligomer sizes, and chose select conformations present at the end of 80 M time units to display in Figure 4.15A-M. The propensity for pore formation monotonically increased with the assembly size, occurring in; 48% of monomers, 65% of dimers, 75% of trimers,75% of tetramers, 83% of pentamers, and 100% of octamers.

Porelike structures which were characterized by the CTR forming a closed loop (CTR-pores) on

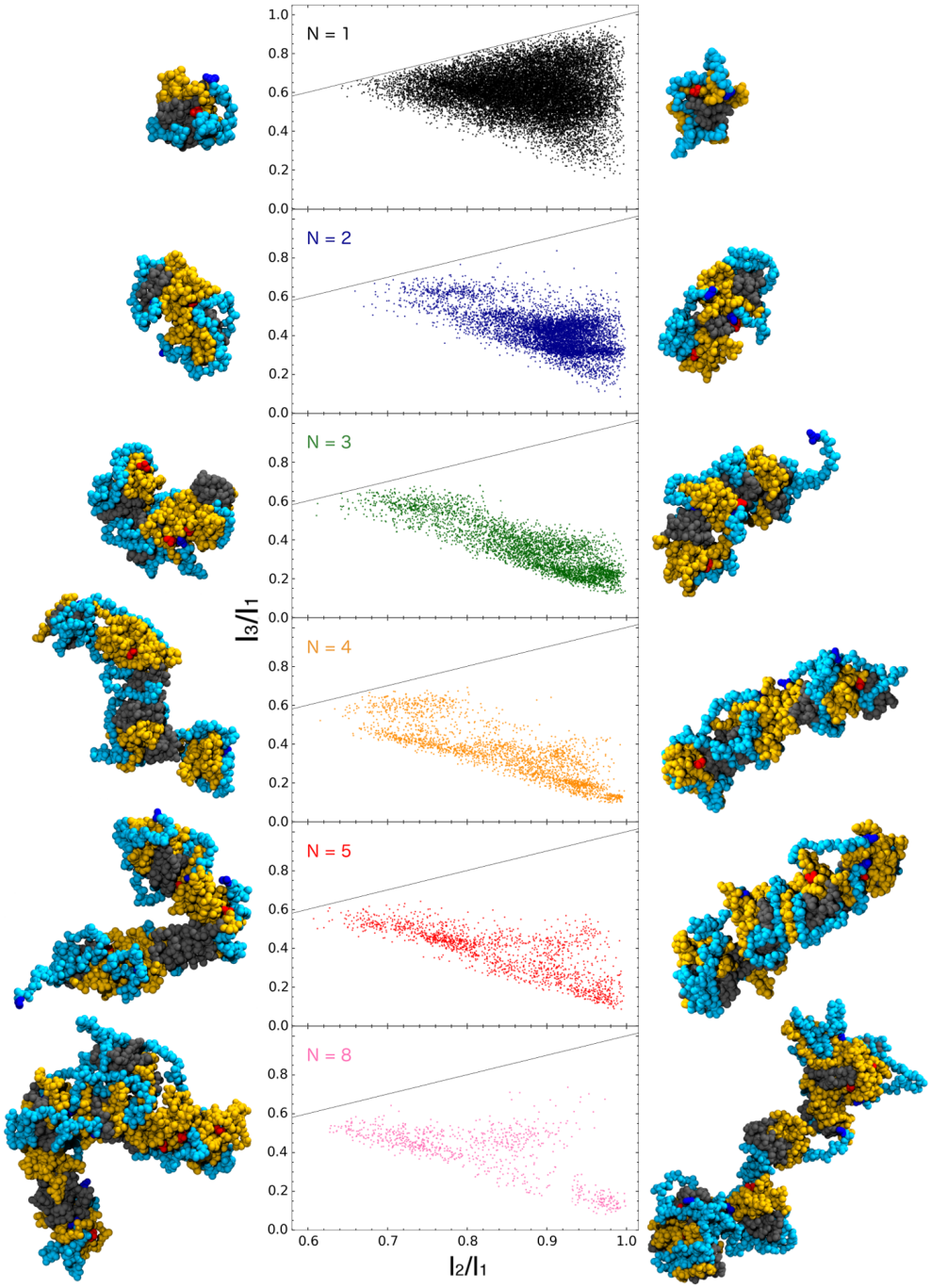

Figure 4.14: αS monomer and oligomer morphologies. 2D distribution of the ratios I_2/I_1 and I_3/I_1, the principal moments of inertia I_2 and I_3 with respect to the largest moment of inertia, I_1. Typical morphologies are displayed on either side of each plot for all assembly sizes, with the C-terminus represented by blue spheres and the N-terminus by red spheres. The NTR (yellow spheres), NAC (grey spheres), and CTR (cyan spheres) are also displayed.

or near the surface of the protein, similar to those observed in monomers (Figre 4.6), were present in all assembly states and were nearly ubiquitous in larger oligomers (Figure 4.15B-F,J-M). Oligomers often had multiple pores formed by CTRs of different peptides (Figure 4.15B,D,J,L,M). The potential for the CTR to make a pore via intermolecular contacts often resulted in large pore diameters. The largest diameters were found in octamers, up to ∼20 Å, although pores with diameters up to ∼16 Å were also present in trimers, tetramers, and pentamers. The depth of these CTR-pores was relatively shallow (∼3-5 Å), as they were most often formed on the surface of the oligomer with the mouth of the pore composed of a single layer of CTR residues.

In addition to the CTR-mediated porelike morphology, we observed other types of porelike morphologies that occurred less frequently and were mediated by heavily NAC and NTR contacts in addition to the CTR. The bent morphologies, which were present in all oligomer orders, feature pores (Figure 4.15A,C,G), with small pore diameters (∼3-6 Å) and depths ∼4-9 Å. Annular structures with a pore in the center were observed in tetramers, pentamers, and octamers (Figure 4.15E,H,K), and were characterized by diameters ∼8-18 Å. Annular pores were largely deeper than CTR-mediated porelike structures, with depths of ∼10-30 Å. The third type of pore was mediated by the branches of "Y"-shaped conformations, which when in contact form a pore. The "Y"-shaped pore morphology is only present in tetramers (Figure 4.15F) and pentamers (Figure 4.14 right side, Figure 4.15I). The "Y"-shaped pore morphology was characterized by a small pore diameter (∼3-6 Å), and depths of ∼5-10 Å.

If these αS porelike oligomers were to get inserted into a membrane and form an ion channel, they could potentially facilitate the flow of ions such as calcium through the membrane and disrupt the homeostasis of the cell. We anticipate that a hydrophobic membrane environment would affect the morphology of αS oligomers, as there is evidence that the NTR undergoes a conversion into a helical structure upon insertion into a membrane[210]. Regardless, the observation of porelike morphologies formed by spatially separated, unstructured peptides prior to membrane insertion provides a noteworthy insight into potentially pathogenic pathways of αS oligomer formation.

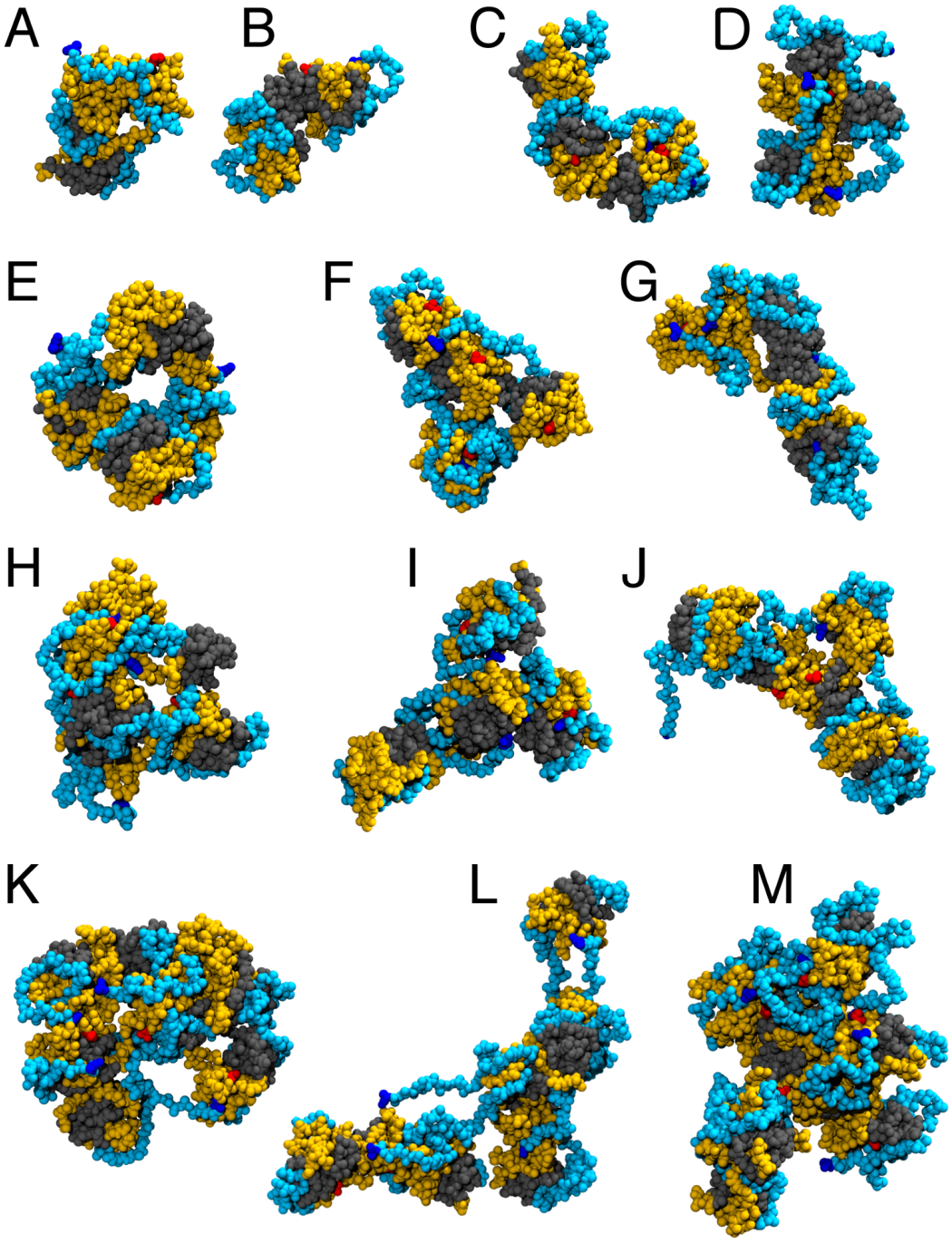

Figure 4.15: Porelike αS conformations. Porelike (A-B) dimer; (C-D) trimer; (E-G) tetramer; (H-J) pentamer; and (K-M) octamer conformations. The C-terminus is represented by blue spheres and the N-terminus by red spheres. The NTR (yellow spheres), NAC (grey spheres), and CTR (cyan spheres) are also displayed.

4.3 Conclusions and Discussion

Studying IDPs such as αS present obvious challenges regardless of the employed methods of study, especially when combined with the relatively lengthy primary structure of the 140 amino acids-long peptide. We have here undertaken, to the best of our knowledge, the first computational study of αS assembly into oligomers from initially separated and unstructured peptides, via DMD4B-HYDRA simulations. We explored a wide range of effective electrostatic interaction strengths and found that the resulting oligomer distribution for $E_{CH} = 0.20$ best matched experimental αS oligomer distribution characteristics measured by PICUP and SDS-PAGE. The DMD4B-HYDRA oligomer distribution at $E_{CH} = 0.20$ was characterized by a monotonically decreasing propensity from monomers to septamers, followed by an increase in octamers. Monomer through pentamer and octamer conformations from simulations with $E_{CH} = 0.20$ were structurally analyzed.

Distributions of the NT-CT distance reaction coordinate revealed two distinct free energy minima which were present in all assembly states. The distinguishing characteristic between these two minima was whether or not M1 and A140 made an intramolecular contact with each other. This feature was absent in conformations from simulations in which E_{CH} was set to 0.00, and thus the electrostatic interaction was essential to the formation of this feature. We examined monomer conformations from these two distinct minima and found significant differences in solvent accessibility, distance from the CM, and secondary structure along the αS sequence. Monomer conformations from both minima displayed porelike morphologies, whereby the essential component of the pore was the CTR, in particular the N103-Q109, D120- S129, and Y133-E139 regions. Conformations in which M1 and A140 were in contact (cluster 1) on average had larger pore diameters and were more likely to have a pore diameter of at least $\sim$3 Åthan conformations in which M1 and A140 were found farther apart (cluster 2).

Analysis of the average secondary structure content revealed that turn content dominated secondary structure in all assembly sizes ($\sim$40-44%), followed by β-strand ($\sim$23-28%) and statistical coil ($\sim$22-25%). The NAC contained more turn content than the NTR and CTR, and the CTR was the most disordered region of the primary structure, consistent with solid-state NMR data[211].

Tertiary structure was dominated by NTR-NTR and NAC-NAC contacts, whereby NTR-NAC and NTR-CTR intramolecular contacts were stronger in monomers than in oligomers. Quaternary structure was dominated by NTR-NTR and NTR-CTR contacts. The NTR-CTR quaternary contacts increased in strength with increasing oligomer order, suggesting that quaternary contacts between these regions play an important role in formation of LMW oligomers. Indeed, analysis of αS oligomer conformations confirmed that the CTR is actively involved in oligomer formation. NAC-NAC quaternary contacts were notably stronger in octamers than in smaller oligomers, although NTR-CTR contacts were still the strongest quaternary contacts in octamers. The NAC region and G68-A78 in particular has been shown to be essential to the formation and structure of αS fibrils, and relevant to the cytotoxicity associated with the peptide[92–94;208]. The strongest group of NAC-NAC quaternary contacts observed in octamers indeed occurred between residues within the A69-A78 region, although DMD4B-HYDRA oligomers that we analyzed here lack the β-sheet structure that is characteristic of these residues within αS fibrils[94]. The observation that the NAC did not contribute significantly to quaternary structure in assembly sizes up to and including pentamers suggest that this region does not play a direct role in the initial aggregation of αS peptides into LMW oligomers, but rather becomes involved further along the aggregation pathway. Instead, the CTR plays a more active role in the formation of LMW oligomers.

The CTR was, as expected, both the most solvent exposed and farthest region from the CM in all assembly states. A140 experienced the largest decrease in SASA (relative to monomers) upon oligomerization, regardless of the assembly size. This residue formed the most numerous strong intermolecular contacts, which could result in A140 becoming buried within the NTR or NAC. The distance of the CTR from the CM increased with oligomer order at a faster rate than the NTR and NAC region up to and including pentamers.

αS monomers were characterized by predominantly globular and quasi-spherical morphologies. Although there is experimental evidence that αS monomers are found in more extended conformations than those observed here[90;195], *in vivo* data suggests that the N-terminal acetylation of αS which occurs in mammalian cells results in conformations which are more compact than those

observed in a buffer[196], which is more aligned with the monomer morphologies we observe. Furthermore, DMD4B-HYRDA conformations are expected to be more compact than more physiologically accurate fully-atomistic representations would be[60;67]. Dimers, trimers, and tetramers were predominantly elongated and rod-like, however, bent and quasi-spherical conformations were also present. Tetramers and pentamers adopted "Y"-shaped morphologies, as well as "L"-shaped and annular morphologies which we also observed in octamers. An important turning point in αS oligomer morphology occurred in pentamers, as elongated morphologies were no longer the most prevalent morphology at the expense of bent, "L" and "Y"-shaped, and annular morphologies. This trend continued in octamers, as elongated octamer conformations were even more rare. Octamers displayed additional "U"-shaped and disk-shaped (oblate spheroid-like) morphologies. This observed conformational polymorphism, or morphological variation, is a trait that has been shown to occur in IDPs, and increases with oligomer order in αS.

The most important observation of the morphology analysis was the discovery of porelike conformations which were present in all assembly states. The propensity to form pores of any kind (with a pore diameter of at least $\sim$3 Å) monotonically increased with assembly size. This trait was also observed in fully-atomistic oligomers of IDPs Aβ40 and Aβ42[67], suggesting that this may be a general characteristic of IDPs in the context of human disease. The CTR was responsible for all observed pore formation in monomers and continued to play a role in oligomer pore formation, particularly in elongated conformations. This CTR behavior is analogous to the behavior of the NTR observed in fully-atomistic Aβ40 and Aβ42 oligomers, which played a similar role in pore formation[67]. All oligomer orders in addition featured pore structures which prominently involved the NAC and NTR, although these conformations were not in dimers and trimers. We observed porelike structures in all of the various tetramers, pentamers, and octamers morphologies, but annular conformations produced some of the largest ($\sim$8-18 Å in diameter) and deepest ($\sim$10-30 Å) pores of any morphology.

The observation of pores in LMW αS oligomers is directly relevant to the pathology of PD. αS and other amyloidogenic proteins have been shown to form annular, ion channel-like structures which,

when embedded into a membrane, are capable of inducing a flow of Ca^{2+} ions through the membrane. This process can potentially lead to cell death and thus inspired the ion channel hypothesis, which states that the primary cytotoxic nature of certain amyloidogenic proteins is due to the formation of annular pores which disturb the homeostasis of cells[83;98–102;104;108;111;113;137;177;209]. The ion channel hypothesis is also consistent with the proposed relationship between calcium regulation and PD[114;115]. The annular tetramer, pentamer, and octamer morphologies reported in this study are ∼60-90 Å in diameter (Figure 4.15E, H, K), large enough to span a membrane, and contain pores which would allow the passage of Ca^{2+} ions (∼2.3 Å in diameter). Furthermore, we observed morphologies which were not annular in nature but contained pores large enough to allow the passage of Ca^{2+} ions, though the hydration shell around Ca^{2+} could impede their passage through pores which are are close in size to these ions[185].

Experiments and molecular dynamics (MD) simulations of αS pores largely involve annular assemblies ranging from pentameric to proto-fibrillar (∼20-40 or more peptides), and these structures have been shown to form both prior to and after insertion into a membrane[83;100–102;104;107;111–113;209]. Tsigelny et. al. reported pentameric and hexameric αS annular pores via combined MD and electron microscopy, with pore diameters ranging from ∼ 20-50Å with an outer ring diameter of ∼90-150 Å[107]. The largest annular pore diameters we observed (∼18 Å) were very near the lower end of this reported range. The outer diameter of the annular pores we observed in tetramers, pentamers, and octamers varied from ∼60-90 Å(Figure 4.15E, H, K) the largest of which again aligns with the lower end of the above reported range. αS was observed via AFM by Ding et. al. to form annular protofibrils which were significantly larger than those in our study, with outer diameters ranging from ∼320-1800 Åand reported pore diameters which were ≤60 Å[101]. Danzer et. al. reported annular oligomers with diameters from ∼400-450 Åwhich mediated an increase in intracellular calcium, consistent with the ion channel hypothesis[83]. Lashuel et. al. used electron microscopy to observe that αS with the PD-associated A53T mutation formed annular protofibrils (∼110 Å) with pore diameters from ∼20-30 Å, which were more similar to the annular pore diameters we observed in the present study[102]. Although these reports of annular oligomers and protofibrils feature structures which are

predominantly larger than the annular morphologies we observed in LMW oligomers, our results suggest that the propensity to form porelike structures increases with assembly size, and we thus expect that simulations of larger oligomers and protofibrils of αS via DMD4B-HYDRA would also yield annular morphologies. It is essential to remember, especially when comparing our results to MD and experimental studies, that the annular structures in this study were derived by DMD4B-HYDRA rather than by fully-atomistic MD. A conversion from the four-bead model to a fully-atomistic one would result in a 'swelling' of the peptide in order to accommodate the additional molecules, which would surely affect both the pore size and the size of annular morphologies as a whole[60;67]. Our study provides further evidence to the notion that porelike αS oligomers can form prior to their insertion into a membrane.

The annular morphologies in our study share characteristics with some models of αS pores, such as a ring-like shape and oligomer order. A tetrameric assembly size was suggested in a model by Tosatto et. al.[111], and a highly structured pentameric pore morphology matched the size and shape of annular structures observed by electron microscopy in a study by Tsigelny et. al.[107]. However, pore models propose a significant amount of ordered structure. β-barrels, which are composed of several β-strand structures in a barrel-shaped morphology, have been suggested as a model for αS pores formed by 20-24 αS peptides[102]. Similarly, the barrel-stave model used by Tosatto et. al. predicts a characteristic α-helical structure[111;212]. Ding et. al. model the formation of annular protofibrils from unfolded αS via the formation of β-sheets[101]. We do not observe this predicted α-helical or β-sheet structure in the annular pore conformations from our study[101;107;111;212]. There is experimental evidence, however, that suggests ion channel-like structures may be significantly more disordered than these models suggest; multiple antimicrobial peptides in a membrane environment were found to form pores which lack the ordered structure predicted by theoretical pore models[178–180]. Combined with the observations in this study, this reasoning suggests that ion channels formed by αS may be similarly disordered. Alternatively, our simulations may not be able to sample more ordered and structured annular conformations that may evolve on longer time scales. Our findings provide important insight into the potential cytotoxicity of αS oligomers within the

context of the ion channel hypothesis[83;101;102;104;107;111–113;209].

4.4 Methods

4.4.1 Simulation protocol

We modeled αS using a four-bead model combined with discrete molecular dynamics and an implicit-solvent force field with amino acid-specific interactions, DMD4B-HYDRA, as described in 2[1;136]. The systems in our simulations were initially composed of 16 spatially separated unfolded αS peptides in a cubic box, with each side 500 nm long and periodic boundary conditions. This results in a concentration of $\sim$3 mM, which is larger than most if not all experimental concentrations of αS, but allows us to reduce the amount of simulation time required for peptides to interact and form oligomers. For each E_{CH} value we examined, 32 replicas of this initial system were created by turning off the force field and performing high temperature simulations (T = 4.0) to generate a series of initial configurations for the production runs at the physiological temperature of 310K (T = 0.130). Each trajectory was 80×10^6 time units, or $\approx 24\mu s$ long with snapshots taken every 0.1×10^6 time units. We simulated seven E_{CH} values: 0.00, 0.01, 0.10, 0.20, 0.30, 0.40, and 0.50.

4.4.2 Structural Analysis

Structural analysis of monomer and oligomer species entailed examining each peptide in a given oligomer individually, unless otherwise noted in the manuscript. For example, for each tetramer conformation we examined, a data point is produced for each of the four peptides which compose the pentamer. Average values of per-oligomer analysis is averaged over each peptide within a given oligomer, rather than being averaged over the entire oligomer. Conducting analysis in this manner allows us to examine the structure of peptides in the context of oligomer conformations.

Potential of the Mean Force

The potential of the mean force (PMF) calculation entailed projecting each conformation onto two selected reaction coordinates; the sum of the solvent accessible surface area of the hydrophobic amino acids in the sequence (according to the Kyte-Doolittle scale[3]) and the distance from the N-terminus of each peptide to the center of mass of each monomer/oligomer. Within the conformational space

of these two reaction coordinates, we created a normalized histogram and counted the number of conformations, N_i, within each bin of the histogram. These histograms were used to create the distributions for these two reaction coordinates. The PMF values were then calculated via the equation $-k_bT \ln(N_i/N_t)$, where N_t is the total number of conformations for a given assembly state and alloform. Every conformation from all 80 M time units from each trajectory was used in the PMF calculation.

Secondary Structure Analysis

Analysis for average secondary structure propensities was calculated as follows: the average secondary structure propensity for each trajectory of a given oligomer size and species is first calculated, both per-residue and averaged over the each peptide individually, using the STRIDE program[157] implemented within the VMD software package[169] to identify secondary structures. These values were then used to calculate a histogram for each secondary structure, whose mean is presented as the average secondary structure propensities in Figure 4.9. The error associated with these averages is the standard error of the mean (SEM) for these histograms, calculated as the corrected sample standard deviation divided by the number of trajectories used for analysis. Monomer and oligomer conformations from the final 20M time units from each trajectory were used for this calculation.

Contact Maps

Two amino acids were considered in contact if the distance between their C_α atoms or C_β- atoms were within a distance of 7.5 Å from each other, consistent with previous DMD and MD studies[20;21;60;69]. The contact maps are presented as an (i,j) matrix with the number of contacts between two specific amino acids averaged over all analyzed conformations. Intramolecular contacts include contact between every (i,j) pair of amino acids within a given peptide (tertiary contacts), and intermolecular contacts include every (i,j) pair of amino acids that belong to different peptides within an oligomer (quaternary contacts). The intramolecular and intermolecular contact maps are presented in a single plot, with intramolecular contacts below the diagonal and intermolecular contacts above the diagonal. Intramolecular contact between an amino acid and any of the two neighboring amino acids on each side are not considered, as their geometry results in these (i,j) pairs falling within the cutoff

distance by default. The color scale corresponds to the propensity for an (i,j) pair of amino acids to be in contact, calculated as the number of observed contacts between an (i,j) pair divided by the number of possible contacts that can occur across the analyzed frames. SEM values were not included as they were negligible when compared to the calculated value due the large number of analyzed conformations. Monomer and oligomer conformations from the final 20M time units from each trajectory were used for this calculation.

Coarse-Grained Solvent Accessible Surface Area

Solvent accessible surface area (SASA) was calculated within the VMD software package[169]. The calculation utilized a spherical surface 1.4 Å away from the surface of each atom in the four-bead model as the cutoff. This calculation was applied to all atoms in an amino acid to give a resulting SASA value. The SASA-per-residue values were attained by averaging over every analyzed conformation. Error bars correspond to SEM values. Monomer and oligomer conformations from the final 20M time units from each trajectory were used for this calculation.

Distance From the Center of Mass

The distance to center of mass for each residue was calculated by first determining the center of mass for each monomer/oligomer via the VMD software package[169]. The location of the C_α atom for each residue is calculated and the distance from this atom to the center of mass of the monomer/oligomer is calculated. The data presented is averaged over every analyzed conformation. Error bars correspond to SEM values. Monomer and oligomer conformations from the final 20M time units from each trajectory were used for this calculation.

Principal Moments of Inertia (Morphology)

The tensor of the moment of inertia was calculated in VMD, allowing us to obtain the three principal moments of inertia for each monomer and oligomer conformation. Monomer and oligomer conformations from the final 20M time units from each trajectory were used for this calculation.

Chapter 5: Final Remarks

In this thesis, we have examined populations of monomers and oligomers of the IDPs Aβ and αS. The computational approaches applied in the studies in this thesis are valuable tools for studying the assembly of proteins, and the study of αS in Chapter 4 is, to our best knowledge, the first such study of αS oligomerization from unfolded and spatially separated peptides. It is crucial to understand the mechanisms and nature of IDP oligomer formation, as LMW oligomeric assemblies may represent the primary toxic agents of certain human diseases. Oligomer formation is an important step in the AD and PD pathology pathway that potential therapies could seek to exploit.

In Chapter 4, the DMD4B-HYDRA approach lead to the observation of annular porelike structures in αS monomers and oligomers. The DMD4B-HYDRA-derived Aβ oligomers which were used to create the fully-atomistic representations in Chapter 3 did not display porelike conformations, however porelike oligomers were observed in the resulting fully-atomistic oligomers of Aβ40 and Aβ42. It would be interesting to explore if porelike conformations persisted or were altered in fully-atomistic αS monomers and oligomers converted from the DMD4B-HYDRA-derived conformations presented in Chapter 4, although the increased size of αS relative to Aβ may present computational challenges. Regardless, the novel observation of porelike assemblies in LMW oligomers of Aβ and αS in this thesis reflect the value of the DMD4B-HYDRA and multi-scale computational approaches. The ion channel hypothesis presents a promising perspective in our understanding of the pathology behind AD and PD. The observation of disordered, annular porelike assemblies formed by LMW oligomers of Aβ and αS in the absence of a membrane from unstructured and spatially divided peptides, and in particular the observation that the propensity for pore formation increased with assembly size in both Aβ and αS, has never been reported before to the best of our knowledge. These results add valuable insights to the ion channel hypotheses of AD and PD.

Bibliography

[1] F. Ding, J. M. Borreguero, S. V. Buldyrev, H. E. Stanley, and N. V. Dokholyan. Mechanism for the α-Helix to β-Hairpin Transition. *Proteins: Struct., Funct., Genet.*, 53(2):220–228, Nov 2003. doi: 10.1002/prot.10468.

[2] Linfeng Wu, Sophie I. Candille, Yoonha Choi, Dan Xie, Lihua Jiang, Jennifer Li-Pook-Than, Hua Tang, and Michael Snyder. Variation and genetic control of protein abundance in humans. *Nature*, 499(7456):79–82, JUL 4 2013.

[3] J. Kyte and R. F. Doolittle. A Simple Method for Displaying the Hydropathic Character of a Protein. *J. Mol. Biol.*, 157(1):105–132, May 1982.

[4] Z. S. Shi, K. Chen, Z. G. Liu, and N. R. Kallenbach. Conformation of the backbone in unfolded proteins. *Chem. Rev.*, 106(5):1877–1897, 2006.

[5] Reinhard Schweitzer-Stenner. Conformational propensities and residual structures in unfolded peptides and proteins. *Mol. BioSyst.*, 8(1):122–133, 2012.

[6] Siobhan E. Toal and Reinhard Schweitzer-Stenner. Local order in the unfolded state: Conformational biases and nearest neighbor interactions. *Biomolecules*, 4(3):725–773, 2014.

[7] Derya Meral, Siobhan Toal, Reinhard Schweitzer-Stenner, and Brigita Urbanc. Water-Centered Interpretation of Intrinsic pPII Propensities of Amino Acid Residues: In Vitro-Driven Molecular Dynamics Study. *J. Phys. Chem. B*, 119(42):13237–13251, OCT 22 2015.

[8] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. The protein data bank. *Nucleic Acids Res.*, 28:235–242, 2000.

[9] C. Levinthal. How to fold graciously. *Mossbauer Spectroscopy In Biological Systems*, pages 22–24, 1969.

[10] F. H. Crick. On protein synthesis. *Symposia of the Society for Experimental Biology*, 12:138, 1958.

[11] Michel Morange. The Protein Side of the Central Dogma: Permanence and Change. *History and Philosophy of the Life Sciences*, 28(4):513–524, 2006.

[12] Christian B. Anfinsen. Principles that Govern the Folding of Protein Chains. *Science*, 181 (4096):223–230, 1973.

[13] J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes. Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins*, 21(3):167–195, 1995.

[14] Vivek Sharma, Ville R. I. Kaila, and Arto Annila. Protein folding as an evolutionary process. *Physica A: Statistical Mechanics and its Applications*, 388(6):851–862, 2009.

[15] Stephen C. Harrison and Richard Durbin. Is There a Single Pathway for the Folding of a Polypeptide Chain? *Proceedings of the National Academy of Sciences of the United States of America*, 82(12):4028–4030, 1985.

[16] Robert A. Goldbeck, Yiren G. Thomas, Eefei Chen, Raymond M. Esquerra, and David S. Kliger. Multiple Pathways on a Protein-Folding Energy Landscape: Kinetic Evidence. *Proceedings of the National Academy of Sciences of the United States of America*, 96(6):2782–2787, 1999.

[17] S. W. Englander and Leland Mayne. The nature of protein folding pathways. *Proceedings of the National Academy of Sciences of the United States of America*, 111(45):15873–15880, 2014.

[18] RJ Ellis and AP Minton. Cell biology - Join the crowd. *Nature*, 425(6953):27–28, SEP 4 2003.

[19] RJ Ellis and AP Minton. Protein aggregation in crowded environments. *Biol. Chem.*, 387(5): 485–497, MAY 2006.

[20] B. Urbanc, L. Cruz, S. Yun, S. V. Buldyrev, G. Bitan, D. B. Teplow, and H. E. Stanley. *In silico* study of amyloid β-protein folding and oligomerization. *Proc. Natl. Acad. Sci. U. S. A.*, 101(50):17345–17350, Dec 2004. doi: 10.1073/pnas.0408153101. URL `http://dx.doi.org/10.1073/pnas.0408153101`.

[21] Brigita Urbanc, Mark Betnel, Luis Cruz, Gal Bitan, and David B. Teplow. Elucidation of amyloid β-protein oligomerization mechanisms: Discrete molecular dynamics study. *J. Am. Chem. Soc.*, 132:4266–4280, 2010.

[22] C. M. Roth, B. L. Neal, and A. M. Lenhoff. Van der Waals interactions involving proteins. *Biophysical Journal*, 70(2):977–987, 1996.

[23] R. K. Pathria. *Statistical mechanics*. Butterworth-Heinemann, Oxford;Boston;, third edition, 2011.

[24] A. K. Dunker, J. D. Lawson, Celeste J. Brown, Ryan M. Williams, Pedro Romero, Jeong S. Oh, Christopher J. Oldfield, Andrew M. Campen, Catherine M. Ratliff, Kerry W. Hipps, Juan Ausio, Mark S. Nissen, Raymond Reeves, ChulHee Kang, Charles R. Kissinger, Robert W. Bailey, Michael D. Griswold, Wah Chiu, Ethan C. Garner, and Zoran Obradovic. Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling*, 19(1):26–59, 2001.

[25] Vladimir N. Uversky, Vrushank Dave, Lilia M. Iakoucheva, Prerna Malaney, Steven J. Metallo, Ravi R. Pathak, and Andreas C. Joerger. Pathological unfoldomics of uncontrolled chaos: intrinsically disordered proteins and human diseases. *Chemical reviews*, 114(13):6844, 2014.

[26] M. M. Babu. The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. *Biochemical Society Transactions*, 44(5):1185–1200, 2016.

[27] Austin Huang and Collin M. Stultz. The Effect of a ΔK280 Mutation on the Unfolded State of a Microtubule-Binding Repeat in Tau. *PLoS Computational Biology*, 4(8), 2008.

[28] Vladimir N. Uversky. Intrinsically disordered proteins from A to Z. *International Journal of Biochemistry and Cell Biology*, 43(8):1090–1103, 2011.

[29] G.B. Irvine, O.M. El-Agnaf, G.M. Shankar, and D.M. Walsh. Protein aggregation in the brain: The molecular basis for Alzheimer's and Parkinson's Diseases. *Mol. Med.*, 14:451–464, 2008.

[30] Robert L. Nussbaum and Christopher E. Ellis. Alzheimer's Disease and Parkinson's Disease. *The New England Journal of Medicine*, 348(14):1356–1364, 2003.

[31] G G Glenner and C W Wong. Alzheimer's disease: Initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochem Biophys Res Commun*, 120:885–890, 1984.

[32] CL Masters, G Simms, NA Weinman, G Multhaup, BL McDonald, and K Beyreuther. Amyloid plaque core protein in Alzheimer-disease and Down syndrome. *Proc. Natl. Acad. Sci. U. S. A.*, 82(12):4245–4249, 1985.

[33] E Levy, M D Carman, I J Fernandez-Madrid, M D Power, I Lieberburg, S G van Duinen, G T A M Bots, W Luyendijk, and B Frangione. Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch-type. *Science*, 248:1124–1126, 1990.

[34] M. Citron, T. Oltersdorf, C. Haass, L. McConlogue, A. Y. Hung, P. Seubert, C. Vigo-Pelfrey, I. Lieberburg, and D. J. Selkoe. Mutation of the β-amyloid precursor protein in familial Alzheimer's disease increases β-protein production. *Nature*, 360(6405):672–674, 1992.

[35] Bart De Strooper, Robert Vassar, and Todd Golde. The secretases: enzymes with therapeutic potential in Alzheimer disease. *Nat. Rev. Neurol.*, 6(2):99–107, FEB 2010.

[36] Christian Haass, Christoph Kaether, Gopal Thinakaran, and Sangram Sisodia. Trafficking and Proteolytic Processing of APP. *Cold Spring Harb. Perspect. Med.*, 2(5):a006270, 2012.

[37] J. T. Jarrett and E. P. Berger and P. T. Lansbury. The C-terminus of the β protein is critical in amyloidogenesis. *Ann N Y Acad Sci*, 695(NIL):144–148, 1993.

[38] J. T. Jarrett, E. P. Berger, and P. T. Lansbury. The Carboxy–Terminus of the β Amyloid Protein is Critical for the Seeding of Amyloid Formation: Implications for the Pathogenesis of Alzheimer's Disease. *Biochemistry*, 32(18):4693–4697, May 1993.

[39] J. A. Hardy and G. A. Higgins. Alzheimer's Disease: The Amyloid Cascade Hypothesis. *Science*, 256:184–185, 1992.

[40] John Hardy and Dennis J. Selkoe. The Amyloid Hypothesis of Alzheimer's Disease: Progress and Problems on the Road to Therapeutics. *Science*, 297(5580):353–356, Jul 2002.

[41] J. Hardy. Alzheimer's disease: Genetic evidence points to a single pathogenesis. *Ann Neuro*, 54(2):143–144, 2003.

[42] Marina D Kirkitadze, Gal Bitan, and David B Teplow. Paradigm Shifts in Alzheimer's Disease and Other Neurodegenerative Disorders: The Emerging Role of Oligomeric Assemblies. *J Neurosci Res*, 69(5):567–577, Sep 2002.

[43] W. L. Klein. ADDLs & protofibrils - the missing links? *Neurobiol. Aging.*, 23:231–233, 2002.

[44] W. L. Klein, W. B. Stine, and D. B. Teplow. Small assemblies of unmodified amyloid β-protein are the proximate neurotoxins in Alzheimer's disease. *Neurobiol Aging*, 25(5):569–580, 2004.

[45] R. Roychaudhuri, M. Yang, M. M. Hoshi, and D. B. Teplow. Amyloid β-protein assembly and Alzheimer disease. *J. Biol. Chem.*, 284:4749–53., 2008.

[46] D. Flöck, S. Colacino, G. Colombo, and A. Di Nola. Misfolding of the amyloid β-protein: A molecular dynamics study. *Proteins*, 62:183–192, 2006.

[47] Alexis Huet and Philippe Derreumaux. Impact of the Mutation A21G (Flemish Variant) on Alzheimer's β-Amyloid Dimers by Molecular Dynamics Simulations. *Biophys. J.*, 91(10): 3829–3840, 2006.

[48] Edgar Luttmann and Gregor Fels. All-atom molecular dynamics studies of the full-length β-amyloid peptides. *Chem. Phys.*, 323:138–147, 2006.

[49] Bogdan Tarus, John E Straub, and D. Thirumalai. Dynamics of Asp23-lys28 salt-bridge formation in Aβ_{10-35} monomers. *J. Am. Chem. Soc.*, 128(50):16159–16168, Dec 2006. doi: 10.1021/ja064872y. URL `http://dx.doi.org/10.1021/ja064872y`.

[50] G. Wei and J.-E. Shea. Effects of solvent on the structure of the Alzheimer amyloid-β(25-35) peptide. *Biophys. J.*, 91:1638–1647, 2006.

[51] Duilio F. Raffa and Arvi Rauk. Molecular Dynamics Study of the β Amyloid Peptide of Alzheimer's Disease and Its Divalent Copper Complexes. *J. Phys. Chem. B*, 111:3789–3799, 2007.

[52] Takako Takeda and Dmitri K. Klimov. Dissociation of Aβ_{16-22} amyloid fibrils probed by molecular dynamics. *J Mol Biol*, 368(4):1202–13, 2007.

[53] Ram Prasad Bora and Rajeev Prabhakar. Translational, rotational and internal dynamics of amyloid β-peptides (Aβ40 and Aβ42) from molecular dynamics simulations. *J. Chem. Phys.*, 131:155103, 2009.

[54] Fu-Feng Liu, Luo Ji, Xiao-Yan Dong, and Yan Sun. Molecular Insight into the Inhibition Effect of Trehalose on the Nucleation and Elongation of Amyloid β-Peptide Oligomers. *J. Phys. Chem. B*, 113(32):11320–11329, AUG 13 2009.

[55] E. P. Raman, T. Takeda, and D. K. Klimov. Molecular Dynamics Simulations of Ibuprofen Binding to Aβ Peptides. *Biophys. J.*, 97:2070–2079, 2009.

[56] C. Wu, M. M. Murray, S. L. Bernstein, M. M. Condron, G. Bitan, J.E. Shea, and M. T. Bowers. The Structure of Aβ42 C-Terminal Fragments Probed by a Combined Experimental and Theoretical Study. *J. Mol. Biol.*, 387:492–501, 2009.

[57] Cao Yang, Xiaolei Zhu, Jinyu Li, and Kai Chen. The effect of solvents on the conformations of amyloid β-peptide (1-42) studied by molecular dynamics simulation. *J. Mol. Struc. - THEOCHEM*, 895:1–8, 2009.

[58] Cao Yang, Xiaolei Zhu, Jinyu Li, and Kai Chen. Molecular dynamics simulation study on conformational behavior of Aβ(1-40) and Aβ(1-42) in water and methanol. *J. Mol. Struc. - THEOCHEM*, 907:51–56, 2009.

[59] Y. Shafrir, S. Durell, A. Anishkin, and R. Guy. Beta-barrel models of soluble amyloid beta oligomers and annular protofibrils. *Proteins: Struct., Funct., Bioinf.*, 78(16):3458–3472, 2010.

[60] Bogdan Barz and Brigita Urbanc. Dimer Formation Enhances Structural Differences between Amyloid β-Protein (1-40) and (1-42): An Explicit-Solvent Molecular Dynamics Study. *PLoS One*, 7:e34345, 2012.

[61] Man Hoang Viet and Mai Suan Li. Amyloid peptide Aβ_{40} inhibits aggregation of Aβ_{42}: Evidence from molecular dynamics simulations. *J. Chem. Phys.*, 136(24), JUN 28 2012.

[62] C. Wu and J. E. Shea. Binding of Congo red to amyloid protofibrils of the Alzheimer Aβ(9-40) peptide probed by molecular dynamics simulations. *Biophys. J.*, 2012. Submitted.

[63] Hyunbum Jang, Laura Connelly, Fernando Teran Arce, Srinivasan Ramachandran, Bruce L. Kagan, Ratnesh Lal, and Ruth Nussinov. Mechanisms for the Insertion of Toxic, Fibril-like β-Amyloid Oligomers into the Membrane. *J. Chem. Theory Comput.*, 9(1):822–833, JAN 2013.

[64] Yiling Xiao, Buyong Ma, Dan McElheny, Sudhakar Parthasarathy, Fei Long, Minako Hoshi, Ruth Nussinov, and Yoshitaka Ishii. Aβ(1-42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease. *Nat. Struct. Mol. Biol.*, 22(6):499–U97, JUN 2015.

[65] Phuong H Nguyen, Fabio Sterpone, Josep M Campanera, Jessica Nasica-Labouze, and Philippe Derreumaux. Impact of the A2V Mutation on the Heterozygous and Homozygous Aβ1–40 Dimer Structures from Atomistic Simulations. *ACS Chem. Neurosci.*, 2016.

[66] Hisashi Okumura and Satoru G. Itoh. Structural and fluctuational difference between two ends of Aβ amyloid fibril: MD simulations predict only one end has open conformations. *Sci. Rep.*, 6, DEC 9 2016.

[67] M. J. Voelker, B. Barz, and B. Urbanc. Fully-atomistic Aβ40 and Aβ42 oligomers in water: Observation of porelike conformations. *J. Chem. Theory Comput.*, 13(9):4567–4583, 2017.

[68] A. Lam, B. Urbanc, J. M. Borreguero, N. D. Lazo, D. B. Teplow, and H. E. Stanley. Discrete molecular dynamics study of Alzheimer amyloid β-protein (Aβ) folding. *Proceedings 2006 International Conference Bioinformatics and Computational Biology (BIOCOMP'06), CSREA Press, Las Vegas, Nevada*, pages 322–328, 2006.

[69] Sijung Yun, Brigita Urbanc, Luis Cruz, Gal Bitan, David B. Teplow, and H. Eugene Stanley. Role of Electrostatic Interactions in Amyloid β-protein (Aβ) Oligomer Formation: A Discrete Molecular Dynamics Study. *Biophys. J.*, 92:4064–4077, Feb 2007. doi: 10.1529/biophysj.106.097766. URL `http://dx.doi.org/10.1529/biophysj.106.097766`.

[70] Brigita Urbanc, Mark Betnel, Luis Cruz, Huiyuan Li, Erica A. Fradinger, Bernhard H. Monien, and Gal Bitan. Structural Basis of Aβ_{1-42} Toxicity Inhibition by Aβ C-Terminal Fragments: Discrete Molecular Dynamics Study. *J. Mol. Biol.*, 410:316–328, 2011.

[71] S. Yun, S. Peng, L. Cruz, S. V. Buldyrev, D. B. Teplow, H. E. Stanley, and B. Urbanc. Effects of the Arctic mutation on oligomer and protofibril formation of Aβ(16-22): A discrete molecular dynamics study. page In preparation, 2011.

[72] Mark Betnel, Nikolay V. Dokholyan, and Brigita Urbanc. *From disordered amyloid β-proteins to soluble oligomers and protofibrils using Discrete Molecular Dynamics, in Alzheimer's disease: Molecular basis of amyloid-β protein aggregation and fibril formation—Insights into low molecular weight and cytotoxic aggregates from computer simulations*, volume 7, chapter 12, pages 333–358. Imperial College Press, London, UK, 2013.

[73] Derya Meral and Brigita Urbanc. Discrete Molecular Dynamics Study of Oligomer Formation by N-Terminally Truncated Amyloid β-Protein. *J. Mol. Biol.*, 425:2260–2275, 2013.

[74] M. Žganec, N. Kruczek, and B. Urbanc. Amino acid substitutions [K16A] and [K28A] distinctly affect amyloid β-protein oligomerization,. *J. Biol. Phys.*, 42:453–476, 2016.

[75] T. L. Tekirian, A. Y. Yang, C. Glabe, and J. W. Geddes. Toxicity of Pyroglutaminated Amyloid β-Peptides 3(pE)-40 and -42 Is Similar to That of Aβ1-40 and -42. *J. Neurochem.*, 73(4):1584–1589, 1999.

[76] Ihsen Youssef, Sabrina Florent-Béchard, Catherine Malaplate-Armand, Violette Koziel, Bernard Bihain, Jean-Luc Olivier, Brigitte Leininger-Muller, Badreddine Kriem, Thierry Oster, and Thierry Pillot. N-truncated amyloid-β oligomers induce learning impairment and neuronal apoptosis. *Neurobiololy of Aging*, 29(9):1319–1333, 2008.

[77] O. M. El-Agnaf, R. Jakes, M. D. Curran, and A. Wallace. Effects of the mutations Ala30 to Pro and Ala53 to Thr on the physical and morphological properties of α-synuclein protein implicated in Parkinson's disease. *FEBS Lett.*, 440:67–70, 1998.

[78] O. M. El-Agnaf, R. Jakes, M. D. Curran, D. Middleton, R. Ingenito, E. Bianchi, A. Pessi, D. Neill, and A. Wallace. Aggregates from mutant and wild-type α-synuclein proteins and NAC peptide induce apoptotic cell death in human neuroblastoma cells by formation of β-sheet and amyloid-like filaments. *FEBS Lett.*, 440:71–75, 1998.

[79] A. B. Singleton, M. Farrer, J. Johnson, A. Singleton, S. Hague, J. Kachergus, M. Hulihan, Peuralinna. T., A. Dutra, R. Nussbaum, S. Lincoln, A. Crawley, M. Hanson, D. Maraganore, C. Adler, M. R. Cookson, M. Muenter, M. Baptista, D. Miller, J. Blancato, J. Hardy, and K. Gwinn-Hardy. α-Synuclein locus triplication causes Parkinson's disease. *Science*, 302:841, 2003.

[80] M.-C. Chartier-Harlin, J. Kachergus, C. Roumier, V. Mouroux, X. Douay, S. Lincoln, C. Levecque, L. Larvor, J. Andrieux, M. Hulihan, N. Waucquier, L. Defebvre, P. Amouyel, M. Farrer, and A. Destée. α-Synuclein locus duplication as a cause of familial Parkinson's disease. *Lancet*, 364:1167–1169, 2004.

[81] P. Ibañez, A. M. Bonnet, B. Débarges, E. Lohmann, Tison. F., P. Pollak, Y. Agid, A. Dürr, and A. Brice. Causal relation between α-synuclein gene duplication and familial Parkinson's disease. *Lancet*, 364:1169–1171, 2004.

[82] E. A. Greenbaum, C. L. Graves, A. J. Mishizen-Eberz, M. A. Lupoli, D. R. Lynch, S. W. Englander, P. H. Axelsen, and B. I. Giasson. The E46K mutation in α-synuclein increases amyloid fibril formation. *J. Biol. Chem.*, 280:7800–7807, 2005.

[83] K. M. Danzer, D. Haasen, A. R. Karow, S. Moussaud, M. Habeck, A. Giese, H. Kretzschmar, B. Hengerer, and M. Kostka. Different species of α-synuclein oligomers induce calcium influx and seeding. *J. Neurosci.*, 27:9220–9232, 2007.

[84] T. F. Outeiro, P. Putcha, J. E. Tetzlaff, R. Spoelgen, M. Koker, F. Carvalho, B. T. Hyman, and P. J. McLean. Formation of toxic oligomeric α-synuclein species in living cell. *PLoS One*, 3:e1867, 2008.

[85] D. P. Karpinar, K. Giller, S. Becker, and M. Baldus. Pre-fibrillar α-synuclein variants with impaired β-structure increase neurotoxicity in Parkinson's disease models. *EMBO J.*, 28(20): 3256–3268, 2009.

[86] Beate Winner, Roberto Jappelli, Samir K. Maji, Paula A. Desplats, Leah Boyer, Stefan Aigner, Claudia Hetzer, Thomas Loher, Marcal Vilar, Silvia Campioni, Christos Tzitzilonis, Alice Soragni, Sebastian Jessberger, Helena Mira, Antonella Consiglio, Emiley Pham, Eliezer Masliah, Fred H. Gage, and Roland Riek. In vivo demonstration that α-synuclein oligomers are toxic. *Proc. Natl. Acad. Sci. U. S. A.*, 108(10):4194–4199, MAR 8 2011.

[87] Lorraine V. Kalia, Suneil K. Kalia, Pamela J. McLean, Andres M. Lozano, and Anthony E. Lang. α-Synuclein oligomers and clinical implications for Parkinson disease. *Ann. Neurol.*, 73 (2):155–169, FEB 2013.

[88] Martin Ingelsson. α-Synuclein oligomers – neurotoxic molecules in Parkinson's disease and other Lewy body disorders. *Front. Neurosci.*, 10, SEP 5 2016.

[89] Kenji Ueda, Hisashi Fukushima, Eliezer Masliah, Yu Xia, Akihiko Iwai, Makoto Yoshimoto, Deborah A. C Otero, Jun Kondo, Yasuo Ihara, and Tsunao Saitoh. Molecular Cloning of cDNA Encoding an Unrecognized Component of Amyloid in Alzheimer's Disease. *Proceedings of the National Academy of Sciences of the United States of America*, 90(23):11282–11286, 1993.

[90] P. H. Weinreb, W. G. Zhen, A. W. Poon, K. A. Conway, and P. T. Lansbury. NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry*, 35(43): 13709–13715, OCT 29 1996.

[91] H.N. Du, L. Tang, X.Y. Luo, H.T. Li, J. Hu, J.W. Zhou, and H.Y. Hu. A peptide motif consisting of glycine, alanine, and valine is required for the fibrillization and cytotoxicity of human α-synuclein. *Biochemistry*, 42:8870–8879, 2003.

[92] A. M. Bodles, D. J. Guthrie, B. Greer, and G. B. Irvine. Identification of the region of non-Aβ component (NAC) of Alzheimer's disease amyloid responsible for its aggregation and toxicity. *J. Neurochem.*, 384-95:384–95, 2001.

[93] B. I. Giasson, I. V. J. Murray, J. Q. Trojanowski, and V. M. Y. Lee. A hydrophobic stretch of 12 amino acid residues in the middle of α-synuclein is essential for filament assembly. *J. Biol. Chem.*, 276(4):2380–2386, JAN 26 2001.

[94] Jose A. Rodriguez, Magdalena I. Ivanova, Michael R. Sawaya, Duilio Cascio, Francis E. Reyes, Dan Shi, Smriti Sangwan, Elizabeth L. Guenther, Lisa M. Johnson, Meng Zhang, Lin Jiang, Mark A. Arbing, Brent L. Nannenga, Johan Hattne, Julian Whitelegge, Aaron S. Brewster, Marc Messerschmidt, Bastien Boutet, Nicholas K. Sauter, Tamir Gonen, and David S. Eisenberg. Structure of the toxic core of α-synuclein from invisible crystals. *Nature*, 525(7570): 486+, SEP 24 2015.

[95] Matthew M. Dedmon, Kresten Lindorff-Larsen, John Christodoulou, Michele Vendruscolo, and Christopher M. Dobson. Mapping long-range interactions in alpha-synuclein using spin-label NMR and ensemble molecular dynamics simulations. *Journal of the American Chemical Society*, 127(2):476–477, 2005.

[96] Tim Bartels, Joanna G. Choi, and Dennis J. Selkoe. α-Synuclein occurs physiologically as a helically folded tetramer that resists aggregation. *Nature*, 477(7362):107–110, 2011.

[97] G. Rossetti, F. Musiani, E. Abad, D. Dibenedetto, H. Mouhib, C. O. Fernandez, and P. Carloni. Conformational ensemble of human α-synuclein physiological form predicted by molecular simulations. *Physical chemistry chemical physics : PCCP*, 18(8):5702–5706, 2016.

[98] N. Arispe, H. B. Pollard, and E. Rojas. β-Amyloid Ca^{2+}-channel hypothesis for neuronal death in Alzheimer-disease. *Mol. Cell Biochem.*, 140(2):119–125, NOV 23 1994. 4th Fall Symposium of the Membrane-Biophysics-Subgroup of the Biophysical-Society, BEAUFORT, NC, OCT, 1993.

[99] Kawahara, M. and Kuroda, Y. and Arispe, N. and Rojas, E. Alzheimer's β-Amyloid, Human Islet Amylin, and Prion Protein Fragment Evoke Intracellular Free Calcium Elevations By a Common Mechanism in a Hypothalamic GnRH Neuronal Cell Line. *J. Biol. Chem.*, 275(19): 14077–14083, 2000.

[100] M. J. Volles, S. J. Lee, J. C. Rochet, M. D. Shtilerman, T. T. Ding, J. C. Kessler, and P. T. Lansbury. Vesicle permeabilization by protofibrillar α-synuclein: Implications for the pathogenesis and treatment of Parkinson's disease. *Biochemistry*, 40(26):7812–7819, JUL 3 2001.

[101] Tomas T. Ding, Seung-Jae Lee, Jean-Christophe Rochet, and Peter T. Lansbury, Jr. Annular alpha-synuclein protofibrils are produced when spherical protofibrils are incubated in solution or bound to brain-derived membranes. *Biochemistry*, 41(32):10209–10217, 2002.

[102] Hilal A. Lashuel, Benjamin M. Petre, Joseph Wall, Martha Simon, Richard J. Nowak, Thomas Walz, and Peter T. Lansbury. α-Synuclein, Especially the Parkinson's Disease-associated Mutants, Forms Pore-like annular and Tubular Protofibrils. *J. Mol. Biol.*, 322(5):1089–1102, 2002.

[103] M.J. Volles and P.T. Jr. Lansbury. Vesicle Permeabilization by Protofibrillar α-Synuclein is Sensitive to Parkinson's Disease-linked Mutations and Occurs by a Pore-like Mechanism. *Biochemistry*, 41:4595–4602, 2002.

[104] H. A. Lashuel, D. Hartley, B. M. Petre, T. Walz, and P. T. Lansbury. Neurodegenerative disease - Amyloid pores from pathogenic mutations. *Nature*, 418(6895):291, JUL 18 2002.

[105] Nelson Arispe, Juan C. Diaz, and Olga Simakova. Aβ ion channels. Prospects for treating Alzheimer's disease with Aβ channel blockers. *Biochim. Biophys. Acta–Biomembranes*, 1768 (8):1952–1965, AUG 2007.

[106] R. Lal, H. Lin, and A. P. Quist. Amyloid β ion channel: 3D structure and relevance to amyloid channel paradigm. *Biochim. Biophys. Acta*, 1768:1966–1975, 2007.

[107] I.F. Tsigelny, P. Bar-On, Y. Sharikov, L. Crews, M. Hashimoto, M.A. Miller, S.H. Keller, O. Platoshyn, J.X. Yuan, and E. Masliah. Dynamics of α-synuclein aggregation and inhibition of pore-like oligomer development by β-synuclein. *FEBS J.*, 274:1862–1877, 2007.

[108] S. M. Butterfield and H. A. Lashuel. Amyloidogenic Protein Membrane Interactions: Mechanistic Insight from Model Systems. *Angew. Chem. Int. Ed. Engl.*, 49:5628–5654, 2010.

[109] Y. Shafrir, S. Durell, N. Arispe, and R. Guy. Models of membrane-bound Alzheimer's Abeta peptide assemblies. *Proteins: Struct., Funct., Bioinf.*, 78:3473–3487, 2010.

[110] T. L. Williams and L. C. Serpell. Membrane and surface interactions of Alzheimer's Aβ peptide – insights into the mechanism of cytotoxicity. *FEBS J.*, 278:3905–3917, 2011.

[111] Laura Tosatto, Alberto O. Andrighetti, Nicoletta Plotegher, Valeria Antonini, Isabella Tessari, Leonardo Ricci, Luigi Bubacco, and Mauro Dalla Serra. Alpha-synuclein pore forming activity upon membrane association. *Biochim. Biophys. Acta–Biomembranes*, 1818(11):2876–2883, NOV 2012.

[112] Igor F. Tsigelny, Yuriy Sharikov, Wolfgang Wrasidlo, Tania Gonzalez, Paula A. Desplats, Leslie Crews, Brian Spencer, and Eliezer Masliah. Role of α-synuclein penetration into the membrane in the mechanisms of oligomer pore formation. *FEBS J.*, 279(6):1000–1013, MAR 2012.

[113] Coralie Di Scala, Nouara Yahi, Sonia Boutemeur, Alessandra Flores, Lea Rodriguez, Henri Chahinian, and Jacques Fantini. Common molecular mechanism of amyloid pore formation by Alzheimer's β-amyloid peptide and α-synuclein. *Sci. Rep.*, 6, JUN 29 2016.

[114] Anthony H. V. Schapira. Calcium dysregulation in Parkinson's disease. *Brain : a journal of neurology*, 136(Pt 7):2015–2016, 2013.

[115] D. J. Surmeier, Paul T. Schumacker, Jaime D. Guzman, Ema Ilijic, Ben Yang, and Enrico Zampese. Calcium and Parkinson's disease. *Biochemical and Biophysical Research Communications*, 483(4):1013–1019, 2017.

[116] Michael J. Berridge. Calcium regulation of neural rhythms, memory and Alzheimer's disease: Calcium regulation of neural rhythms. *The Journal of Physiology*, 592(2):281–293, 2014.

[117] N. Arispe, H. B. Pollard, and E. Rojas. Giant multilevel cation channels formed by Alzheimer disease amyloid β-protein [AβP-(1-40)] in bilayer membranes. *Proc. Natl. Acad. Sci. U. S. A.*, 90:10573–10577, 1993.

[118] M. Kawahara and Y. Kuroda. Molecular Mechanism of Neurodegeneration Induced By Alzheimer's β-Amyloid Protein: Channel Formation and Disruption of Calcium Homeostasis. *Brain Research Bulletin*, 53(4):389–397, 2000.

[119] S. R. Durell, G. R. Guy, N. Arispe, R. Eduardo, and H. B. Pollard. Theoretical models of the ion channel structure of amyloid β-protein. *Biophys. J.*, 67:2137–2145, 1994.

[120] H. Jang, F. T. Arce, S. Ramachandran, R. Capone, R. Lal, and R. Nussinov. β-Barrel Topology of Alzheimer's β-Amyloid Ion Channels. *J. Mol. Biol.*, 404(5):917–934, 2010.

[121] L. Connelly, F. T. Arce, H. Jang, C. Ricardo, S.A. Kotler, S. Ramachandran, B. L. Kagan, R. Nussinov, and R. Lal. Atomic force microscopy and MD simulations reveal pore-like structures of all-D-Enantiomer of Alzheimer's β-Amyloid Peptide: Relevance to the ion channel mechanism of AD pathology. *J. Phys. Chem. B*, 116(5):1728–1735, 2012.

[122] Charles K. Fisher and Collin M. Stultz. Constructing ensembles for intrinsically disordered proteins. *Current Opinion in Structural Biology*, 21(3):426–431, 2011.

[123] H. J. C. Berendsen, D. van der Spoel, and R. van Drunen. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.*, 91(1-3):43–56, September 1995.

[124] Erik Lindahl, Berk Hess, and David van der Spoel. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Model.*, 7(8):306–317, 2001. URL `http://dx.doi.org/10.1007/s008940100045`.

[125] David Van Der Spoel, Erik Lindahl, Berk Hess, Gerrit Groenhof, Alan E. Mark, and Herman J. C. Berendsen. GROMACS: fast, flexible, and free. *J. Comput. Chem.*, 26(16):1701–1718, 2005. ISSN 0192-8651. doi: 10.1002/jcc.20291. URL `http://onlinelibrary.wiley.com/doi/10.1002/jcc.20291/abstract`.

[126] Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, 4(3):435–447, March 2008. doi: 10.1021/ct700301q. URL `http://dx.doi.org/10.1021/ct700301q`.

[127] William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79(2):926–935, 1983. ISSN 00219606. doi: 10.1063/1.445869. URL `http://link.aip.org/link/JCPSA6/v79/i2/p926/s1&Agg=doi`.

[128] William L. Jorgensen and Julian Tirado-Rives. The OPLS Potential Functions for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.*, 110 (6):1657–1666, 1988.

[129] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.*, 118(45):11225–11236, January 1996. doi: 10.1021/ja9621760. URL `http://dx.doi.org/10.1021/ja9621760`.

[130] George A. Kaminski, Richard A. Friesner, Julian Tirado-Rives, and William L. Jorgensen. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B*, 105(28):6474–6487, July 2001.

[131] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, 8:3684–3690, 1984.

[132] Olgun Guvench and Alexander D. Jr. MacKerell. Comparison of protein force fields for molecular dynamics simulations. *Methods Mol. Biol.*, 443:63–88, 2008.

[133] B. Hess and N. F. A. van der Vegt. Hydration Thermodynamic Properties of Amino Acid Analogues: A Systematic Comparison of Biomolecular Force Fields and Water Models. *J. Phys. Chem. B*, 110:17616–17626, 2006.

[134] Joao Henriques, Carolina Cragnell, and Marie Skepo. Molecular Dynamics Simulations of Intrinsically Disordered Proteins: Force Field Evaluation and Comparison with Experiment. *J. Chem. Theory Comput.*, 11(7):3420–3431, JUL 2015.

[135] Kresten Lindorff-Larsen, Stefano Piana, Ron O. Dror, and David E. Shaw. How Fast-Folding Proteins Fold. *Science*, 334(6055):517–520, 2011.

[136] D. C. Rapaport. *The art of molecular dynamics simulation.* Cambridge University Press, Cambridge, 1997.

[137] H. B. Pollard, E. Rojas, and N. Arispe. A new hypothesis for the mechanism of amyloid toxicity, based on the calcium channel activity of amyloid β protein (AβP) in phospholipid bilayer membranes. In R. M. Nitsch, J. H. Growdon, S. Corkin, and R. J. Wurtman, editors, *Alzheimer's disease: Amyloid precusor proteins, signal transduction, and neuronal transplantation*, volume 695 of *Ann. N. Y. Acad. Sci.*, pages 165–168. 1993.

[138] M. Žganec, E. Žerovnik, and B. Urbanc. Self-assembly of globular protein stefin B into polymorphic oligomers probed by discrete molecular dynamics. *J. Chem. Theory Comput.*, 11: 2355–2366, 2015.

[139] Bogdan Barz, Bradley S. Turner, Rama Bansil, and Brigita Urbanc. Folding of Pig Gastric Mucin Non-glycosylated Domains: A Discrete Molecular Dynamics Study. *J. Biol. Phys.*, 38: 681–703, 2012.

[140] T. Creighton. *Proteins: structures and molecular properties, second edition.* W. H. Freeman & Co., New York, 1993.

[141] A. Lam, D. B. Teplow, H. E. Stanley, and B. Urbanc. Effects of the Arctic (E22→G) Mutation on Amyloid β-Protein Folding: Discrete Molecular Dynamics Study. *J. Am. Chem. Soc.*, 130: 17413–17422, 2008.

[142] Shuting Zhang, Dillion M. Fox, and Brigita Urbanc. Insights into Formation and Structure of Aβ Oligomers Cross-Linked via Tyrosines. *J. Phys. Chem. B*, 121(22):5523–5535, JUN 8 2017.

[143] Brigita Urbanc, Jose M. Borreguero, Luis Cruz, and H. Eugene Stanley. *Ab initio* Discrete Molecular Dynamics Approach to Protein Folding and Aggregation. *Methods. Enzymol.*, 412: 314–338, 2006.

[144] D. J. Selkoe. Alzheimer's Disease: Genes, Proteins, and Therapy. *Physiological Reviews*, 81 (2):741–766, Apr 2001.

[145] N Sawamura, M Morishima-Kawashima, H Waki, K Kobayashi, T Kuramochi, M P Frosch, K Ding, M Ito, T W Kim, R E Tanzi, F Oyama, T Tabira, S Ando, and Y Ihara. Mutant Presenilin 2 Transgenic Mice. A large increase in the levels of Aβ42 is presumably associated with the low density membrane domain that contains decreased levels of glycerophospholipids and sphingomyelin. *J. Biol. Chem.*, 275(36):27901–27908, 2000.

[146] K. N. Dahlgren, A. M. Manelli, W. B. Stine, L. K. Baker, G. A. Krafft, and M. J. LaDu. Oligomeric and Fibrillar Species of Amyloid-β Peptides Differentially Affect Neuronal Viability. *J. Biol. Chem.*, 277(35):32046–32053, 2002.

[147] P. Maiti, A. Lomakin, G.B. Benedek, and G. Bitan. Despite its role in assembly, methionine 35 is not necessary for amyloid β-protein toxicity. *J. Neurochem.*, 113:1252–1262, 2010.

[148] Eileen McGowan, Fiona Pickford, Jungsu Kim, Luisa Onstead, Jason Eriksen, Cindy Yu, Lisa Skipper, M. Paul Murphy, Jenny Beard, Pritam Das, Karen Jansen, Michael Delucia, Wen-Lang Lin, Georgia Dolios, Rong Wang, Christopher B Eckman, Dennis W Dickson, Mike Hutton, John Hardy, and Todd Golde. Aβ42 is essential for parenchymal and vascular amyloid deposition in mice. *Neuron*, 47(2):191–199, Jul 2005.

[149] K. Iijima, H.P. Liu, A.S. Chiang, S.A. Hearn, M. Konsolaki, and Y. Zhong. Dissecting the pathological effects of human Aβ40 and Aβ42 in Drosophila: A potential model for Alzheimer's disease. *Proc. Natl. Acad. Sci. U. S. A.*, 101:6623–6628, 2004.

[150] G. Bitan, M. D. Kirkitadze, A. Lomakin, S. S. Vollers, G. B. Benedek, and D. B. Teplow. Amyloid β-protein (Aβ) assembly: Aβ40 and Aβ42 oligomerize through distinct pathways. *Proc. Natl. Acad. Sci. U. S. A.*, 100(1):330–335, 2003.

[151] S. L. Bernstein, N. F. Dupuis, N. D. Lazo, T. Wyttenbach, M. M. Condron, G. Bitan, D. B. Teplow, J.-E. Shea, B. T. Ruotolo, C. V. Robinson, and M. T. Bowers. Amyloid-β protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease. *Nat. Chem.*, 1:326–331, 2009.

[152] T. L. Williams, L. C. Serpell, and B. Urbanc. Stabilization of Native Amyloid β-Protein Oligomers by Copper and Hydrogen Peroxide Induced Cross-Linking of Unmodified Proteins (CHICUP). *Biochim. Biophys. Acta*, 1864:249–259, 2016.

[153] Brigita Urbanc, Luis Cruz, David B. Teplow, and H. Eugene Stanley. Computer Simulations of Alzheimer's Amyloid β-Protein Folding and Assembly. *Curr. Alzheimer Res.*, 3:493–504, 2006.

[154] J.-E. Shea and B. Urbanc. Insights into Aβ aggregation: A molecular dynamics perspective. *Curr. Topics Med. Chem., Special Issue: Protein Misfolding in Conformational Disorders, Guest Ed. Cláudio M. Gomes*, 12:2596–2610, 2012.

[155] A. Samiotakis, D. Homouz, and M.S. Cheung. Multiscale investigation of chemical interference in proteins. *J. Chem. Phys.*, 132:175101, 2010.

[156] Derya Meral and Brigita Urbanc. Erratum to Discrete Molecular Dynamics Study of Oligomer Formation by N-Terminally Truncated Amyloid β-Protein. *J. Mol. Biol.*, 427:2726–2729, 2015.

[157] D. Frishman and P. Argos. Knowledge-Based Protein Secondary Structure Assignment. *Proteins: Struct., Funct., Genet.*, 23:566–579, 1995.

[158] N. G. Sgourakis, Y. Yan, S. A. McCallum, C. Wang, and A. E. Garcia. The Alzheimer's peptides Aβ40 and 42 adopt distinct conformations in water: A combined MD/NMR Study. *J. Mol. Biol.*, 368:1448–1457, 2007.

[159] K. Aurelia Ball, Aaron H. Phillips, David E. Wemmer, and Teresa Head-Gordon. Differences in β-strand Populations of Monomeric Aβ40 and Aβ42. *Biophys. J.*, 104(12):2714–2724, JUN 18 2013.

[160] M. Yang and D. B. Teplow. Amyloid β-Protein Monomer Folding: Free-Energy Surfaces Reveal Alloform-Specific Differences. *J. Mol. Biol.*, 384:450–464, 2008.

[161] Noel D Lazo, Marianne A Grant, Margaret C Condron, Alan C Rigby, and David B Teplow. On the nucleation of amyloid β-protein monomer folding. *Protein Sci*, 14(6):1581–1596, Jun 2005. PMCID:PMC2253382.

[162] Jose M Borreguero, Brigita Urbanc, Noel D Lazo, Sergey V Buldyrev, David B Teplow, and H. Eugene Stanley. Folding events in the 21-30 region of amyloid β-protein (Aβ) studied *in silico*. *Proc. Natl. Acad. Sci. U. S. A.*, 102(17):6015–6020, Apr 2005. doi: 10.1073/pnas.0502006102. URL `http://dx.doi.org/10.1073/pnas.0502006102`.

[163] Aneta T Petkova, Yoshitaka Ishii, John J Balbach, Oleg N Antzutkin, Richard D Leapman, Frank Delaglio, and Robert Tycko. A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Natl. Acad. Sci. U. S. A.*, 99(26): 16742–16747, Dec 2002.

[164] Robert Tycko. Amyloid Polymorphism: Structural Basis and Neurobiological Relevance. *Neuron*, 86(3):632–645, MAY 6 2015.

[165] Luis Cruz, Brigita Urbanc, Jose M Borreguero, Noel D Lazo, David B Teplow, and H. Eugene Stanley. Solvent and mutation effects on the nucleation of amyloid β-protein folding. *Proc. Natl. Acad. Sci. U. S. A.*, 102(51):18258–18263, Dec 2005. PMCID:PMC1317965.

[166] Yilin Yan and Chunyu Wang. Aβ42 is More Rigid Than Aβ40 at the C Terminus: Implications for Aβ Aggregation and Toxicity. *J. Mol. Biol.*, 364:853–862, Sep 2006.

[167] T Luhrs, C Ritter, M Adrian, D Riek-Loher, B Bohrmann, H Dobeli, D Schubert, and R Riek. 3D structure of Alzheimer's amyloid-β(1-42) fibrils. *Proc. Natl. Acad. Sci. U. S. A.*, 102: 17342–17347, 2005.

[168] Karin Rybka, Siobhan E. Toal, Daniel J. Verbaro, Daniel Mathieu, Harald Schwalbe, and Reinhard Schweitzer-Stenner. Disorder and order in unfolded and disordered peptides and proteins: A view derived from tripeptide conformational analysis. II. Tripeptides with short side chains populating asx and β-type like turn conformations. *Proteins: Struct., Funct., Bioinf.*, 81(6):968–983, JUN 2013.

[169] W. Humphrey, A. Dalke, and K. Schulten. VMD: visual molecular dynamics. *J. Mol. Graphics*, 14(1):33–38, Feb 1996.

[170] B. Urbanc. Flexible N-Termini of Amyloid β-Protein Oligomers: A Link between Structure and Activity? *Israel J. Chem.*, 57:651–664, 2017.

[171] S. K. Rhee, A. P. Quist, and R. Lal. Amyloid β protein-(1-42) forms calcium-permeable, Zn^2-sensitive channels. *J. Biol. Chem.*, 273(22):13379–13382, 1998.

[172] J. C. Diaz, J. Linnehan, H. Pollard, and N. Arispe. Histidines 13 and 14 in the Aβ sequence are targets for inhibition of Alzheimer's disease Aβ ion channel and cytotoxicity. *Biol. Res.*, 39(3):447–460, 2006.

[173] S. A. Kotler, P. Walsh, J. R. Brender, and A. Ramamoorthy. Differences between amyloid-β aggregation in solution and on the membrane: Insights towards elucidation of the mechanistic details of Alzheimer's disease. *Chem. Soc. Rev.*, 43(19):6692–6700, 2014.

[174] Ball, K. Aurelia and Phillips, Aaron H. and Nerenberg, Paul S. and Fawzi, Nicolas L. and Wemmer, David E. and Head-Gordon, Teresa. Homogeneous and Heterogeneous Tertiary Structure Ensembles of Amyloid-β Peptides. *Biochemistry*, 50(35):7612–7628, 2011. ISSN 0006-2960.

[175] Yu-Shan Lin, Gregory R. Bowman, Kyle A. Beauchamp, and Vijay S. Pande. Investigating How Peptide Length and a Pathogenic Mutation Modify the Structural Ensemble of Amyloid β Monomer. *Biophys. J.*, 102:315–324, 2012.

[176] B. Strödel, J. W. Lee, C. S. Whittleston, and D. J. Wales. Transmembrane Structures for Alzheimer's Aβ_{1-42} Oligomers. *J. Am. Chem. Soc.*, 132:13300–13312, 2010.

[177] A. Quist, I. Doudevski, H. Lin, R. Azimova, D. Ng, B. Frangione, B. Kagan, J. Ghiso, and R. Lal. Amyloid ion channels: A common structural link for protein-misfolding disease. *Proc. Natl. Acad. Sci. U. S. A.*, 102:10427–10432, 2005.

[178] H. Leontiadou, A. E. Mark, and S. J. Marrink. Antimicrobial peptides in action. *J. Am. Chem. Soc.*, 128(37):12156–12161, 2006.

[179] D. Sengupta, H. Leontiadou, A. E. Mark, and S. J. Marrink. Toroidal pores formed by antimicrobial peptides show significant disorder. *Biochim. Biophys. Acta–Biomembranes*, 1778 (10):2308–2317, 2008.

[180] S. J. Marrink, A. H. de Vries, and D. P. Tieleman. Lipids on the move: Simulations of membrane pores, domains, stalks and curves. *Biochim. Biophys. Acta–Biomembranes*, 1788 (1):149–168, 2009.

[181] K. Ono, M. M. Condron, and D. B. Teplow. Structure-neurotoxicity relationships of amyloid β-protein oligomers. *Proc. Natl. Acad. Sci. U. S. A.*, 106:14745–14750, 2009.

[182] M. Townsend, G. M. Shankar, T. Mehta, D. M. Walsh, and D. J. Selkoe. Effects of secreted oligomers of amyloid β-protein on hippocampal synaptic plasticity: a potent role for trimers. *J. Physiol.*, 572(2):477–492, 2006.

[183] Mathew A. Sherman, Michael LaCroix, Fatou Amar, Megan E. Larson, Colleen Forster, Adriano Aguzzi, David A. Bennett, Martin Ramsden, and Sylvain E. Lesn'e. Soluble Conformers of Aβ and Tau Alter Selective Proteins Governing Axonal Transport. *J. Neurosci.*, 36(37): 9647–9658, SEP 14 2016.

[184] Summer L Bernstein, Thomas Wyttenbach, Andrij Baumketner, Joan-Emma Shea, Gal Bitan, David B Teplow, and Michael T Bowers. Amyloid β-Protein: Monomer Structure and Early Aggregation States of Aβ42 and Its Pro19 Alloform. *J. Am. Chem. Soc.*, 127(7):2075–2084, Feb 2005.

[185] Franca M. Floris, Maurizio Persico, Alessandro Tani, and Jacopo Tomasi. Hydration shell structure of the calcium ion from simulations with ab initio effective pair potentials. *Chemical Physics Letters*, 227(1):126–132, 1994.

[186] Gal Bitan, Sabrina S Vollers, and David B Teplow. Elucidation of Primary Structure Elements Controlling Early Amyloid β-Protein Oligomerization. *J. Biol. Chem.*, 278(37):34882–34889, 2003. PMID:12840029.

[187] G. Bussi, D. Donadio, and M. Parrinello. Canonical sampling through velocity rescaling. *J. Chem. Phys.*, 126:014101, 2007.

[188] M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.*, 52(12):7182, 1981.

[189] R Jakes, M. G. Spillantini, and M. Goedert. Identification of 2 distinct synucleins from human brain. *FEBS Lett.*, 345(1):27–32, MAY 23 1994.

[190] M. G. Spillantini, M. L. Schmidt, V. M. Y. Lee, J. Q. Trojanowski, R. Jakes, and M. Goedert. α-synuclein in Lewy bodies. *Nature*, 388(6645):839–840, AUG 28 1997.

[191] S. Turnbull, B. J. Tabner, O. M. A. El-Agnaf, S. Moore, Y. Davies, and D. Allsop. α-Synuclein implicated in Parkinson's disease catalyses the formation of hydrogen peroxide in vitro. *Free Rad. Biol. Med.*, 30(10):1163–1170, MAY 15 2001.

[192] Sreeganga Chandra, Gilbert Gallardo, Rafael Fernandez-Chacon, Oliver M. Schluter, and Thomas C. Sudhof. α-Synuclein cooperates with CSPα in preventing neurodegeneration. *Cell*, 123(3):383–396, 2005.

[193] Kristin E. Larsen, Yvonne Schmitz, Matthew D. Troyer, Eugene Mosharov, Paula Dietrich, Abrar Z. Quazi, Magali Savalle, Venu Nemani, Farrukh A. Chaudhry, Robert H. Edwards, Leonidas Stefanis, and David Sulzer. α-Synuclein overexpression in PC12 and chromaffin cells impairs catecholamine release by interfering with a late step in exocytosis. *J. Neurosci.*, 26 (46):11915–11922, NOV 15 2006.

[194] David C. DeWitt and Elizabeth Rhoades. αS-Synuclein can inhibit SNARE-mediated vesicle fusion through direct interactions with lipid bilayers. *Biochemistry*, 52(14):2385, 2013.

[195] Bruno Fauvet, Martial K. Mbefo, Mohamed-Bilal Fares, Carole Desobry, Sarah Michael, Mustafa T. Ardah, Elpida Tsika, Philippe Coune, Michel Prudent, Niels Lion, David Eliezer, Darren J. Moore, Bernard Schneider, Patrick Aebischer, Omar M. El-Agnaf, Eliezer Masliah, and Hilal A. Lashuel. α-Synuclein in central nervous system and from erythrocytes, mammalian cells, and Escherichia coli exists predominantly as disordered monomer. *J. Biol. Chem.*, 287 (19):15345–15364, 2012.

[196] Francois-Xavier Theillet, Andres Binolfi, Beata Bekei, Andrea Martorana, Honor May Rose, Marchel Stuiver, Silvia Verzini, Dorothea Lorenz, Marleen van Rossum, Daniella Goldfarb, and Philipp Selenko. Structural disorder of monomeric α-synuclein persists in mammalian cells. *Nature*, 530(7588):45+, FEB 4 2016.

[197] Wei Wang, Iva Perovic, Johnathan Chittuluru, Alice Kaganovich, Linh T. T. Nguyen, Jingling Liao, Jared R. Auclair, Derrick Johnson, Anuradha Landeru, Alana K. Simorellis, Shulin Ju, Mark R. Cookson, Francisco J. Asturias, Jeffrey N. Agar, Brian N. Webb, ChulHee Kang, Dagmar Ringe, Gregory A. Petsko, Thomas C. Pochapsky, and Quyen Q. Hoang. A soluble α-synuclein construct forms a dynamic tetramer. *Proc. Natl. Acad. Sci. U. S. A.*, 108(43): 17797–17802, OCT 25 2011.

[198] Eduardo Coelho-Cerqueira, Phelippe Carmo-Goncalves, Anderson Pinheiro, Juliana Cortines, and Cristian Follmer. α-Synuclein as an intrinsically disordered monomer – fact or artefact? *FEBS J.*, 280(19):4915–4927, 2013.

[199] Kenjiro Ono, Tokuhei Ikeda, Jun-ichi Takasaki, and Masahito Yamada. Familial Parkinson disease mutations influence α-synuclein assembly. *Neurobiol. Dis.*, 43(3):715–724, SEP 2011.

[200] Kenjiro Ono, Hideki Mochizuki, Tokuhei Ikeda, Tomoko Nihira, Jun-ichi Takasaki, David B. Teplow, and Masahito Yamada. Effect of melatonin on α-synuclein self-assembly and cytotoxicity. *Neurobiol. Aging*, 33(9):2172–2185, SEP 2012.

[201] Kenjiro Ono, Jun-ichi Takasaki, Ryoichi Takahashi, Tokuhei Ikeda, and Masahito Yamada. Effects of antiparkinsonian agents on β-amyloid and α-synuclein oligomer formation in vitro. *J. Neurosci. Res.*, 91(10):1371–1381, OCT 2013.

[202] Claudio D. Borsarelli, Lisandro J. Falomir-Lockhart, Veronika Ostatni, Jonathan A. Fauerbach, He-Hsuan Hsiao, Henning Urlaub, Emil Paleek, Elizabeth A. Jares-Erijman, and Thomas M. Jovin. Biophysical properties and cellular toxicity of covalent crosslinked oligomers of α-synuclein formed by photoinduced side-chain tyrosyl radicals. *Free Radic. Biol. Med.*, 53 (4):1004–1015, 2012.

[203] Dhiman Ghosh, Mrityunjoy Mondal, Ganesh M. Mohite, Pradeep K. Singh, Priyatosh Ranjan, A. Anoop, Saikat Ghosh, Narendra Nath Jha, Ashutosh Kumar, and Samir K. Maji. The Parkinson's disease-associated H50Q mutation accelerates α-synuclein aggregation *in vitro*. *Biochemistry*, 52:6925–6927, 2013.

[204] Srabasti Acharya, Brian M. Safaie, Piriya Wongkongkathep, Magdalena I. Ivanova, Aida Attar, Frank-Gerrit Klarner, Thomas Schrader, Joseph A. Loo, Gal Bitan, and Lisa J. Lapidus. Molecular basis for preventing α-synuclein aggregation by a molecular tweezer. *J. Biol. Chem.*, 289(15):10727–10737, 2014.

[205] Jane R. Allison, Peter Varnai, Christopher M. Dobson, and Michele Vendruscolo. Determination of the free energy landscape of alpha-synuclein using spin label nuclear magnetic resonance measurements. *J. Am. Chem. Soc.*, 131(51):18314–18326, 2009.

[206] Diániel Sz"ollősi, Tamás Horváth, Kyou-Hoon Han, Nikolay V. Dokholyan, P'eter Tompa, Lajos Kalmi'ar, and Tam'as Hegedűs. Discrete molecular dynamics can predict helical prestructured motifs in disordered proteins. *PLoS One*, 9(4), APR 24 2014.

[207] Ulf Dettmer, Andrew J. Newman, Eric S. Luth, Tim Bartels, and Dennis Selkoe. In vivo cross-linking reveals principally oligomeric forms of α-synuclein and β-synuclein in neurons and non-neural cells. *J. Biol. Chem.*, 288(9):6371, 2013.

[208] H. N. Du, L. Tang, X. Y. Luo, H. T. Li, J. Hu, J. W. Zhou, and H. Y. Hu. A peptide motif consisting of glycine, alanine, and valine is required for the fibrillization and cytotoxicity of human α-synuclein. *Biochemistry*, 42:8870–8879, 2003.

[209] M. J. Volles and P. T. Jr. Lansbury. Zeroing in on the pathogenic form of α-synuclein and its mechanism of neurotoxicity in Parkinson's disease. *Biochemistry*, 42:7871–7878, 2003.

[210] R. Jr. Bussell and D. Eliezer. A structural and functional role for 11-mer repeats in α-synuclein and other exchangeable lipid binding proteins. *J. Mol. Biol.*, 13:763–778, 2003.

[211] Henrike Heise, Wolfgang Hoyer, Stefan Becker, Ovidiu C. Andronesi, Dietmar Riedel, Marc Baldus, and Adriaan Bax. Molecular-Level Secondary Structure, Polymorphism, and Dynamics of Full-Length α-Synuclein Fibrils Studied by Solid-State NMR. *Proceedings of the National Academy of Sciences of the United States of America*, 102(44):15871–15876, 2005.

[212] L. Yang, T. A. Harroun, T. M. Weiss, L Ding, and H. W. Huang. Barrel-stave model or toroidal model? A case study on melittin pores. *Biophys. J.*, 81(3):1475–1485, SEP 2001.

Appendix A: Supporting Information for: *Fully-Atomistic Aβ40 and Aβ42 Oligomers in Water: Observation of Porelike Oligomers*

A.1 Supporting Information

A.1.1 Detailed Region-Specific Differences in Secondary Structure Between Aβ40 and Aβ42 Conformations

Turn and β-strand propensities for each amino acid in the sequence are displayed in Figure 5 of the main manuscript. In comparison to Aβ40 monomers, Aβ42 monomers had significantly increased turn propensity in the A2-F4 and G37-V40 regions. Relative to Aβ40, increased turn propensity at A2 was observed in all Aβ42 oligomers. In addition, elevated turn propensity was noted in E2 of Aβ42 trimers. Aβ42 pentamers were characterized by increased turn in the D7-S8 region. The most notable increase in turn propensity in Aβ42 relative to Aβ40 involved the MHR and CTR regions (G37-V40 in monomers, V36-V40 in dimers and trimers, only L34 in tetramers, and L34-V40 with exception of G37 in pentamers). The turn propensity in Aβ40 trimers was lower than or equal to that of Aβ42 trimers along the entire sequence. Aβ40 monomers, dimers, and tetramers had higher turn propensity that the respective Aβ42 conformations at R5 and in the D7-S8 region. In addition, Aβ40 monomers had increased turn propensity at G9, Q15, L17, and in the S26-N27 and G33-M35 regions. In comparison to Aβ42 dimers, the E11-V12 region had an increased turn propensity in Aβ40 dimers. Aβ40 tetramers had higher turn propensity than Aβ42 tetramers at E11 and S26. Interestingly, Aβ40 pentamers had increased turn relative to Aβ42 pentamers only in the A30-I31 region, which was unique to pentamers.

Aβ40 and Aβ42 monomers had consistently high turn propensity in the D23-N27 region, as well as a lower but still significant turn propensity in the G33-G37 region. were present in Aβ42 oligomers as well. The turn propensity in the S26-N27 region was the highest in Aβ42 pentamers.

Of all assembly states, Aβ40 and Aβ42 monomers were characterized by the most distinct β-strand propensities. Aβ40 monomers had higher β-strand propensities than Aβ42 monomers in the A2-F4 and G29-I31 regions. Aβ42 monomers, in contrast, has higher β-strand propensities than

Aβ40 monomers in the G9-V12, CHC, G25-S26, I32-M35, and G38-V40 regions. The regions F19-F20, L34-M35, and G38-V39 had the highest β-strand propensities in Aβ42 monomers (15-17%), significantly higher than in Aβ40 monomers. Aβ40 dimers had a higher β-strand propensity than Aβ42 dimers only at A2, whereas the β-strand propensity in Aβ42 dimers was higher than in Aβ40 dimers at R5 as well as in the I31-I32 and V39-V40 regions. Aβ40 trimers had a higher β-strand propensity than Aβ42 trimers in the A2-E3 region, whereas Aβ42 trimers had increased β-strand propensity relative to Aβ40 trimers at E11, N27, I32, and in the V39-V40 region. In Aβ40 tetramers, increased β-strand propensity relative to that of Aβ42 tetramers was noted at E3, Q15, and L34. In contrast, Aβ42 tetramers showed higher β-strand propensity than Aβ40 tetramers only in the C-terminal G38-V40 region. The alloform-specific β-strand propensity differences in the A2-F4 region disappeared in pentamers. Aβ40 pentamers had higher β-strand propensity than Aβ42 pentamers only in the S26-N27 region, whereas Aβ42 pentamers showed higher β-strand propensity than Aβ40 pentamers in the V36-V40 region. A third region in Aβ42 with increased β-strand arose at G29-I32 for Aβ42 in dimers and persisted in trimers, tetramers, and pentamers. The appearance of this third region resulted in a slight shift of the peaks in the β-strand propensity per amino acid from I32-M35 and G38-I41 in Aβ42 monomers to L34-V36 and V39-I41 in Aβ42 oligomers. In Aβ42 pentamers, the these two regions with increased β-strand merged into a single region. In Aβ40 oligomers, distinct peaks in β-strand propensity per amino acid were also observed at the C-terminus. While centered within the regions A30-I32 and M35-V36 in Aβ40 monomers, the respective β-strand propensities increased in Aβ40 dimers and trimers. Although these two regions with increased β-strand were present in Aβ40 tetramers, a third, less pronounced peak in the β-strand propensity in the G38-V39 region of Aβ40 pentamers was observed. Overall, Aβ42 had significantly higher β-strand content than Aβ40 in the CTR not only in monomers but also in all oligomers under study.

Coil propensities per amino acid are shown in Figure 6 of the main manuscript (left panels). Relative to Aβ42 conformations, Aβ40 monomers and oligomers had higher coil propensities at A2 and in the G37-V40 region. Aβ42 monomers had a higher coil propensity than Aβ40 monomers only at G9, Q15, and N27. In addition to the regions mentioned above, Aβ40 dimers had higher

coil propensity than Aβ42 dimers also at V18. The D7-S8 region, V12, and G25 had more coil in Aβ42 dimers than in Aβ40 dimers. In comparison to Aβ42 trimers, we observed higher coil propensity in Aβ40 trimers also in the L34-M35 region. Similarly, V24 also showed increased coil content in Aβ40 relative to Aβ42 tetramers. Aβ42 tetramers had higher coil propensity than Aβ40 tetramers at R5 and in the D7-G9 region. The D7-S8 region had higher coil propensity in Aβ40 pentamers than in Aβ42 pentamers, in contrast to the results for dimers and tetramers. Aβ40 pentamers had increased coil relative to Aβ42 pentamers also at V12 and V36. Aβ42 pentamers were characterized by higher coil propensity than Aβ40 pentamers at L17 and in the A30-I32 region, a feature that was unique to pentamers. The bridge propensities per amino acid are displayed in Figure 6 of the main manuscript (right panels). While these propensities were relatively high, the regions with high bridge propensities roughly matched with hydrophobic regions and there were more residue-to-residue fluctuations than in other secondary structure propensities. Notably, the CTR had higher bridge propensity in Aβ42 than in Aβ40 conformations. At the N-terminus, F4 had a higher bridge propensity in Aβ40 than in Aβ42 conformations, suggesting that the N-terminus was more structured in Aβ40 than in Aβ42.

A.1.2 Detailed Description of Tertiary Structure of Aβ40 and Aβ42 Conformations

In the following description of tertiary structure, we first comment on the strongest tertiary contacts, which occurred close to the diagonal (diagonal contacts), and then proceed to more distant contacts (off-diagonal contacts). Here we follow the same definition of contact map regions as defined in the main manuscript, with the predominant off-diagonal tertiary contacts being enclosed in three frames in Figure 7 of the main manuscript (black frames). Region 1 comprises off-diagonal tertiary contacts of the CHC with the CFR, MHR, and CTR (Figure 7, main manuscript), black trapezoid-like frames). Region 2 contains off-diagonal tertiary contacts of the MHR with CTR (Figure 7, main manuscript), black triangle-like frames). Region 3 corresponds to tertiary contacts of the NTR with the Q15-V40 region (Figure 7, main manuscript), black rectangular frames).

Tertiary Structure of Aβ40 and Aβ42 Monomers

The strongest diagonal tertiary contacts in Aβ40 monomers were in the N-terminal region: H6-G9, H13-K16, H14-L17; in the CFR: A21-V24, D23-S26, G25-N27; and in the MHR: I31-L34. Aβ42 monomers showed, in addition to those tertiary contacts found in Aβ40 monomers, strong contact between D1-F4 and S26-G29. Moreover, the contact I31-L34 found in Aβ40 monomers was shifted by four residues toward the C-terminus, to M35-G38, in Aβ42 monomers. These same diagonal contacts in the C_β–C_β contact maps were notably weaker with the exception of H14-L17, A21-V24, and D23-S26 in Aβ40 monomers, and A21-V24 and D23-S26 for Aβ42 monomers (Figure S2).

The off-diagonal tertiary contacts were dominated by the contacts in Region 1 (Figure 7, main manuscript, black trapezoid-like frames), specifically between V18-A21 and A30-I32 in Aβ40 monomers and between F19-F20 and I32-M35 regions in Aβ42 monomers. C_β–C_β contact maps were very similar, with the addition of prominent contacts at L17-I32 and A21-N27 in Region 1 (Figure S2, black trapezoid-like frames) of Aβ40 monomers and V18-I31 in Region 1 of Aβ42 monomers. In addition, we observed strong tertiary contacts in Region 2 (Figure 7, main manuscript, black triangle-like frames), between I32-V36 and V39-I41, which were stronger in Aβ42 monomers than in Aβ40 monomers. C_β–C_β contact maps showed little variation from these observations, though the contact L34-V40 in Aβ40 became more prominent (Figure S2, black triangle-like frames). Tertiary contacts in Region 3 (Figure 7, main manuscript, black rectangular frames) were distinctly different in Aβ40 and Aβ42 monomers. In Aβ40 monomers, the NTR formed strong contacts with the CHC (with the strongest contact between F4-F19) as well as with the I31-G37 region (with the strongest contact between F4-M35), In Aβ42 monomers, these contacts were significantly weaker. Similar results were observed in Region 3 of the C_β–C_β contact maps (Figure S2, black rectangular frames), with the addition of a prominent contact, F4-I31, in Aβ40 monomers. The group of tertiary contacts between the NTR and the I31-G37 region in Aβ40 monomers extended out of Region 3 and included tertiary contacts between regions S8-Y10 and M35-G37. Additional alloform-specific off-diagonal tertiary contacts were observed between Regions 2 and 3, whereby the tertiary contact V12-V36 that was the strongest in this region in Aβ42 monomers was significantly weaker in Aβ40 monomers. This

contact was also prominent in the C_β–C_β contact maps (Figure S2).

Diagonal Tertiary Contacts in Aβ40 and Aβ42 Oligomers

Diagonal tertiary contacts in Aβ40 oligomers were overall comparable to those in Aβ40 monomers with some variability in strength. Notably, the tertiary contact E3-H6, which was not observed in Aβ40 monomers, characterized all Aβ40 oligomers (Figure 7, main manuscript). This contact was only prominent in Aβ40 tetramers when examining C_β–C_β contact maps (Figure S2). Only Aβ40 tetramers were also characterized by a strong contact R5-S8, a feature not observed in C_β–C_β contact maps. In Aβ40 tetramers and pentamers, the contact I31-L34 shifted to A30-G33 and became less pronounced than in smaller Aβ40 oligomers. In Aβ40 pentamers, the tertiary contacts H13-K16 and H14-L17 were the strongest and the tertiary contact A21-V24 was weaker than in Aβ40 tetramers. Similarly, diagonal tertiary contacts in Aβ42 oligomers were comparable to those in Aβ42 monomers, again with some variability in strength. Interestingly, the tertiary contact V36-V39, which was not present in Aβ42 monomers, characterized Aβ42 dimers. In addition to the diagonal tertiary contacts in Aβ42 conformations described above, a new tertiary contact I31-L34 becomes prominent in C_α–C_α contact maps of Aβ42 trimers, tetramers, and pentamers, but not in C_β–C_β contact maps. Additional diagonal tertiary contacts which were unique to Aβ42 pentamers were observed: D7-Y10 at the N-terminus, G25-K28 and V24-N27 within the CFR and V39-A42 within the CTR, whereas the tertiary contacts H14-L17, A21-V24, and N27-A30 were significantly weaker than in Aβ42 tetramers.

Diagonal contacts were generally much weaker in C_β–C_β contact maps (Figure S2). C_β–C_β contacts between H14-L17 and A21-V24 were notable in Aβ40 dimers, though none were observed in Aβ42 dimers or in Aβ40 trimers. A21-V24 was the only prominent contact in Aβ42 trimers, and was also observed in Aβ42 tetramers. H13-K16, H14-L17, and D23-S26 were prominent contacts in Aβ40 pentamers, while D23-S26 and V36-V39 were also strong diagonal contacts in Aβ42 pentamers.

Tertiary Contacts in Aβ40 and Aβ42 Oligomers: Region 1

Region 1 contacts in Aβ40 trimers were somewhat stronger than in Aβ40 dimers and were dominated by two strong contacts, V18-L34 and F19-L34. Contacts in Region 1 were even stronger in

Aβ40 tetramers and were dominated by L17-I32, L17-G33, V18-M35, and V18-V36. The dominant contacts in Region 1 within Aβ40 pentamers were between resides V18-I32 and between residues F19-I31, which enclosed the CFR. In Aβ42 oligomers, stronger tertiary contacts in Region 1 were shifted toward the CTR with the oligomer order. Of all Aβ40 and Aβ42 oligomers, Aβ42 pentamers had the strongest tertiary contacts in Region 1, dominated by F19-A30, F19-I31, A21-A30, and A21-I31, which enclosed the CFR.

Region 1 of the C_β–C_β contact maps (Figure S2, black trapezoid-like frames) featured similar overall observation to C_α–C_α contact maps, with a few differences in the strong contacts present in Aβ40 and Aβ42 oligomers when compared to Region 1 in C_α–C_α contact maps. Aβ40 dimers feature prominent C_β–C_β contacts which were not present in C_α–C_α contact maps between residues V18-V36 and F19-V36, as did Aβ42 dimers between residues V18-L34 and F19-I32. Aβ40 trimers displayed a strong contact F19-I34, and Aβ42 tetramers had a strong contact between residues F19-V39 which shifted to a strong F19-V40 contact in Aβ42 pentamers.

Tertiary Contacts in Aβ40 and Aβ42 Oligomers: Region 2

Aβ40 dimers and tetramers exhibited off-diagonal tertiary contacts in Region 2 that were comparable to those in Aβ40 monomers (Figure 7, main manuscript, black triangle-like frames). Aβ40 trimers showed weaker contacts and Aβ40 pentamers displayed the weakest contacts in this region. Of all Aβ42 oligomers, Aβ42 trimers had the strongest tertiary contacts in Region 2, dominated by contacts between L34-M35 and V40-I41. This group of contacts extended out of Region 2 and included the CTR of Aβ42. Of all Aβ42 conformations, Aβ42 pentamers had the fewest strong contacts in Region 2 (with I31-V36 as the strongest contact).

Region 2 of the C_β–C_β contact maps (Figure S2, black triangle-like frames) were extremely similar to C_α–C_α contact maps, as Aβ42 trimers similarly having the strongest tertiary contacts. A pair of contacts between Region 1 and 2 were noted in C_β–C_β contact maps of Aβ40 tetramers which were absent from C_α–C_α contact maps, between residues V24-M35 and V24-V36.

Tertiary Contacts in Aβ40 and Aβ42 Oligomers: Region 3

Region 3 covers tertiary contacts between the NTR and remaining peptide regions. The most numerous and the strongest tertiary contacts (between C_α as well as C_β atoms) in Region 3 were observed for Aβ40 and Aβ42 monomers with the former showing stronger contacts that the latter (Figures 7, main manuscript and S2, black rectangular frames). Across all Aβ42 conformations, Region 3 had the weakest C_α–C_α and C_β–C_β tertiary contacts (Figures 7, main manuscript and S2, black rectangular frames), as would be expected based on the predominantly hydrophilic nature of the NTR. In Aβ40 oligomer conformations, however, Region 3 showed stronger tertiary contacts than Region 2, with an exception of the I31-V36 contact in Region 3, which was the strongest contact in Regions 2 and 3. We observed fewer C_α–C_α and C_β–C_β tertiary contacts between the NTR and remaining peptide regions in Aβ40 pentamers than in smaller Aβ40 oligomers (Figures 7, main manuscript and S2, black rectangular frames). In Aβ40 pentamers, there were fewer C_α–C_α tertiary contacts between the NTR and remaining peptide regions than in smaller Aβ40 oligomers. Aβ40 and Aβ42 pentamers had a comparable number and strength of tertiary contacts in Region 3, dominated by contacts between E3-F4 and L34-V36 regions (Aβ40) or F4-R5 and G38-V39 regions (Aβ42). Examination of tertiary C_β–C_β contacts (Figure S2, black rectangular frames) resulted in similar observation.

In-between Regions 1 and 3 (Figure 7, main manuscript, black trapezoid-like and rectangular frames, respectively), we observed unusually strong off-diagonal C_α–C_α tertiary contacts between the V12-Q15 and I32-L34 regions that were specific to Aβ40 pentamers (but not to Aβ42 pentamers).

A.1.3 Detailed Description of Quaternary Structure of Aβ40 and Aβ42 Conformations

To facilitate a comprehensive description of quaternary contacts, we defined three regions that were involved the most in quaternary contact formation (Figure 7, main manuscript, white frames). Region 4 enclosed quaternary CHC-CHC, CHC-MHR, and CHC-CTR contacts (Figure 7, main manuscript, white trapezoid-like frames). Region 5 comprises quaternary MHR-MHR, MHR-CTR, and CTR-CTR contacts (Figure 7, main manuscript, white triangle-like frames). Region 6 covers

quaternary contacts between the NTR and the Q15-V40 region and is thus analogous to Region 3 as defined for tertiary contact maps (Figure 7, main manuscript, white rectangular frames).

Quaternary Contacts in Aβ40 and Aβ42 Oligomers: Region 4

In Aβ40 dimers, the strongest C_α–C_α quaternary contacts in Region 4 were between V18-F19 and G33-L34 (Figure 7, main manuscript, trapezoid-like white frames). Relative to Aβ40 dimers, Aβ40 trimers displayed more C_α–C_α quaternary contacts in Region 4, with the strongest ones between the L17-V18 and L34-M35 regions. C_α–C_α quaternary contacts in Region 4 became more numerous and stronger in Aβ40 tetramers, where two groups of contacts emerged as the strongest: between V18-F19 and L34-V36 as well as between F20-A21 and G29-I31.

Together, these two groups of C_α–C_α quaternary contacts represent an anti-parallel contact interface between pairs of Aβ40 peptides. This anti-parallel quaternary structure feature was even more pronounced in Aβ40 pentamers, which showed an overall a smaller number of C_α–C_α quaternary contacts in Region 4 that were nonetheless significantly stronger than in Aβ40 tetramers. The predominant C_α–C_α quaternary contacts in Aβ40 pentamers in Region 4 were: L17-I32, V18-I31, V18-I32, F19-I32, F20-G29, and A21-G29. Relative to Aβ40 dimers and trimers, another group of C_α–C_α quaternary contacts centered at V18-F19 and F19-F19 became more prominent in Aβ40 tetramers. This group of contacts became even more prominent in Aβ40 pentamers, with the strongest contacts: V18-F20, F19-F19, F19-F20, and F20-F20.

Aβ40 C_β–C_β quaternary contacts in Region 4 (Figure S2, white trapezoid-like frames) were similar to C_α–C_α contacts for Aβ40 dimers and tetramers, though for Aβ40 trimers the strongest contacts were located adjacent to those in the C_α–C_α contact maps, between the V18-F19 and I31-I32 regions. Aβ40 pentamers differed the most between C_β–C_β and C_α–C_α contact maps; with contacts between residues F19-F20, F20-I31, and F19-L34 being much stronger in the C_β–C_β contact maps. The anti-parallel quaternary structure feature present in C_α–C_α contact map of Aβ40 tetramers and pentamers is less prominent in tetramers and not visible in pentamers in the C_β–C_β contact maps.

Interestingly, Aβ42 oligomers were characterized by weaker quaternary contacts in Region 4 than the corresponding Aβ40 oligomers. In Aβ42 dimers, contacts V18-I32 and V18-G33 dominated the

quaternary structure in Region 4. Aβ42 trimers had more contacts in Region 4 but these were overall weaker than in Aβ42 dimers. Notably, quaternary contacts F19-F19 and F19-F20 were the strongest in Aβ42 trimers. Unlike for Aβ40 tetramers and pentamers, we did not observe any anti-parallel motif in Region 4 for Aβ42 tetramers and pentamers. Aβ42 tetramers had a larger number of strong contacts in Region 4 than Aβ42 trimers, with the strongest contacts between V18 and the G33-I41 region. Quaternary contacts F19-F19 and V18-F19 were more prominent in Aβ42 tetramers than in Aβ42 trimers. Aβ42 pentamers exhibited strong quaternary contacts: L17-F19, V18-V18, V18-F19, V18-F20, and F19-F19, which dominated Region 4. The other contacts in Aβ42 tetramers were fewer and shifted to the CTR, whereby the strongest contacts were between F19-F20 and V39. Aβ42 C_β–C_β quaternary contacts in Region 4 (Figure S2, white trapezoid-like frames) were similar to Aβ42 C_α–C_α quaternary contacts for all oligomers. A minute difference was observed in Aβ42 trimers, which had a stronger contact between F19-F19 in the C_β–C_β contact map.

Quaternary Contacts in Aβ40 and Aβ42 Oligomers: Region 5

The number and strength of quaternary contacts in Region 5 increased with oligomer order in Aβ40 conformations with the strongest quaternary contacts in Aβ40 pentamers (Figure 7, main manuscript, white triangle-like frames). Whereas Aβ40 dimers had no strong quaternary contacts in Region 5, two relatively strong contacts were found adjacent to Region 5: A30-I31 and A30-I32. In Aβ40 trimers, I32-V36 emerged as the strongest contact in Region 5. In Aβ40 tetramers, the strongest contact in Region 5 was V36-V36, followed by L34-V36. Unlike Aβ40 trimers, Aβ40 tetramers were characterized with relatively strong diagonal quaternary contacts involving the M35-G37 region. Additional albeit not very strong contacts emerged in Aβ40 tetramers between the MHR and CTR. The predominant quaternary contacts in Region 5 in Aβ40 pentamers were between the MHR regions with the strongest contact L34-V36. Of all Aβ40 oligomers, the diagonal quaternary contacts in Region 5 were the strongest in Aβ40 pentamers, dominated by L34-L34, M35-M35, and V39-V39.

The number and strength of quaternary contacts in Region 5 increased from Aβ42 dimers through Aβ42 tetramers (Figure 7, main manuscript white triangle-like frames). There were no strong

quaternary contacts in Aβ42 dimers. In Aβ42 trimers, the strongest contact was I32-I32 and the neighboring contacts that were adjacent to Region 5. These contacts became much weaker in Aβ42 tetramers, which instead displayed a strong contact G37-G38 and additional diagonal contacts G37-G37 and G38-G38. Quaternary contacts between MHR and V40 were also notable in Aβ42 tetramers. The strongest quaternary contacts in this region in Aβ42 pentamers were I31-V36, I31-G37, and neighboring contacts adjacent to Region 5 that included A30-G37. The diagonal quaternary contacts that became more prominent in Aβ42 pentamers (relative to smaller Aβ42 oligomers) were observed between pairs of the I32-V39 regions, dominated by V39-V39, G38-G38, and L34-L34. Relatively strong diagonal contacts adjacent to Region 5 were G29-G29 and A30-A30.

Region 5 of C_β–C_β contact maps (Figure S2, white triangle-like frames) were overwhelmingly similar to C_α–C_α contacts for dimers and trimers of both alloforms. We noted that for tetramers, several C_β–C_β contacts in Region 5 were stronger than their corresponding C_α–C_α contacts: I31-I31, L34-L34, and M35-M35 for Aβ40 and M35-V36 and V36-V36 for Aβ42. These stronger diagonal contacts in Aβ40 tetramers further hint at merging parallel cross-β structure. Similarly for pentamers, I31-I31, V36-V36, L34-V36, and V36-V39 for Aβ40 and I31-I32, I31-M35, I32-M35, V36-V36, L34-V39, and V36-I41 (just outside of Region 5) for Aβ42 were stronger in the C_β–C_β contact maps versus the C_α–C_α contact maps. Conversely, for Aβ42 pentamers, the A30-G37 and V39-V39 C_β–C_β contacts were weaker than their C_α–C_α counterparts.

Quaternary Contacts in Aβ40 and Aβ42 Oligomers: Region 6

Overall, of Regions 4-6, Region 6 had the lowest density of quaternary contacts (Figure 7, main manuscript, white rectangular frames). The predominant quaternary contacts in Region 6 in Aβ40 dimers were F4-A30 and F4-V39, whereas in Aβ40 trimers A2-V36, A2-G37, and A2-F20 were the strongest. Aβ40 tetramers had the fewest contacts in Region 6, dominated by D1-V40, A2-V40, and F4-G37, however, they were characterized by unique albeit relatively weak quaternary contacts joining Regions 4 and 6. Of all Aβ40 and Aβ42 oligomers, Aβ40 pentamers had the most and the strongest contacts in Region 6. In Aβ40 pentamers, quaternary contacts extend out of Region 6 toward Region 4. We noted three groups of contacts between regions: (a) F4-H6 and F20-E22 (with

the strongest contacts R5-E22, H6-A21, and H6-E22) — this group extended out of Region 6 and included also contacts G9-E22 and G9-E23; (b) F4-H6 and MHR; and (c) Y10-V12 and I31-I32 (outside but adjacent to Region 6) — these contacts which were connected to a distinct group of parallel contacts between regions G9-V12 and N27-A30.

Additional quaternary contacts not included in Region 6 were the NTR-NTR contacts that contributed to quaternary structure of Aβ40 trimers and to a lesser extent Aβ40 tetramers, but were absent from Aβ40 dimers and pentamers. Aβ42 dimers and trimers had no strong contacts in Region 6. In Aβ42 trimers, a group of quaternary contacts between Regions 4 and 6 centered around E11-G33 was observed. Aβ42 tetramers displayed two larger groups of contacts in Region 6 centered around the strongest contacts A2-A21 and F4-G33 and another group of contacts centered at Y10-I32 between Regions 4 and 6. The overall strongest contact in Region 6, D1-D23, was observed only in Aβ42 pentamers. This contact was peculiar because it corresponded to a proximity of two negatively charged aspartic acids. The corresponding quaternary structure based on proximity between C_β atoms exerted the same strong contact D1-D23 (Figure S2). Aβ42 pentamers displayed quaternary contacts between the NTR and CFR, mostly between D1-A2 and E22-V24 regions, which included the contact D1-D23. Additional contacts in Region 6 in Aβ42 pentamers were: F4-K16, F4-L17, and F4-L34. We also noted numerous contacts in Aβ42 pentamers between Regions 4 and 6, dominated by V12-F20 and V12-A21. Outside Region 6, notable NTR-NTR contacts in Aβ42 pentamers were F4-R5 and R5-R5. The corresponding quaternary contacts between C_β atoms were also observed but were somewhat weaker (Figure S2).

Region 6 of C_β–C_β contact maps (Figure S2, white rectangular frames) was similar to C_α–C_α contact maps for dimers of both alloforms and Aβ42 trimers. For Aβ40 trimers, the C_β–C_β contact F4-M35 was more prominent than its C_α–C_α counterpart, while D1-A2 (outside of Region 6) was less strong. Aβ40 tetramers did not clearly have the fewest C_β–C_β contacts in Region 6, unlike in C_α–C_α contact maps. The region was otherwise very similar between the C_β–C_β and C_α–C_α contact maps of Aβ40 tetramers, with the exception of contact D1-V40, which was notably weaker in the C_β–C_β contact maps. Aβ42 C_β–C_β and C_α–C_α contact maps were also very similar for

tetramers, with the exception of F4-L34 being notably more prominent in the C_β–C_β contact maps. Aβ40 pentamers had the most differences when comparing C_β and C_α–C_α contact maps. When comparing the three regions mentioned above for Aβ40 pentamers: (a) was overwhelmingly similar, (b) was notably weaker in the C_β–C_β contact maps than in the C_α–C_α contact maps, and (c) has notably stronger contacts between residues Y10-I31, V12-I31 and V12-I32 in the C_β–C_β contact maps. C_β–C_β contact V12-V12, between Region 1 and Region 2 was also notably stronger than the corresponding C_α contact for Aβ40 pentamers. The notable D1-D23 contact was also present in C_β–C_β contact maps of Aβ42 pentamers. Contacts between F4-F20 and F4-A21 were notably more prominent in C_β–C_β contact maps of Aβ42 pentamers than in C_α–C_α contact maps, while F4- R5 and R5-R5 (outside of Region 6) were weaker. Other observations made above for the Aβ42 pentamer C_α–C_α contact map were extremely similar when we examined the corresponding C_β–C_β contact map.

A.1.4 Detailed Distance from the Center of Mass (CM) Analysis

Aβ40 monomers, dimers, and trimers exhibited larger distances than the corresponding Aβ42 conformations in the CTR for monomers and dimers or at V40 for trimers (Figure S3, left panels). Aβ40 tetramers were characterized by larger distances from the CM than Aβ42 tetramers at D1 and in the V39-V40 region. Aβ42 tetramers displayed larger distances from the CM than Aβ40 tetramers in the G9-H13, A21-V24, and K28-V36 regions. Pentamers had the largest variation in distances between the two different Aβ species. Aβ40 pentamers exhibited larger distances than Aβ42 pentamers at D1 and in the H14-V18 and M35-V40 regions. In contrast, Aβ42 pentamers showed increased distances relative to Aβ40 pentamers in the G9-V12 and A21-G29 region, the latter almost exactly coinciding with the CFR.

The distance from the CM per residue for each alloform individually was plotted to elucidate how the amino acid arrangement with respect to the CM changed with oligomer order (Figure S3, left panels). Overall, distances increased with the oligomer order for both alloforms. In Aβ40 and Aβ42 monomers, the L17-F19 region was the closest and D1 was the furthest to the CoM. Aβ40 dimers and trimers were characterized by two regions V18-F19 and I31-I32 that were the most proximate

to the CoM. Whereas D1 remained the furthest from the CM in Aβ40 dimers, in Aβ40 trimers, the region D7-S8 was further away from the CM than D1. In Aβ40 tetramers, the region L34-V36 became the closest to the CoM, while D1 and the D7-S8 region were the most distant. The amino acids that were the closest to and the furthest from the CM in Aβ40 pentamers were F20 and D1, respectively. In Aβ42 dimers, there were several amino acids that showed the smallest distance from the CoM: V18-F19, I32, and V39-V40, whereas D1 remained the most distant from the CoM. Aβ42 trimers showed similar characteristics with two regions, V18-F19 and I32-L34, that were the most proximate to the CoM. whereas D1 as well as the D7-S8 region were the furthest from the CoM. In Aβ42 tetramers, F19 and the V39-V40 region were the most proximate to the CM whereas the D7-G9 region became the most distant from the CoM. Relative to smaller Aβ42 oligomers, Aβ42 pentamers showed a shift of the regions that were the most proximate to the CM toward the C-terminus (the V36-G37 and V39-V40 regions). The peptide region that was the furthest away from the CM in Aβ42 pentamers was D1-H14 with S8-G9 at the largest distance from the CoM.

A.1.5 Detailed Alloform- and Assembly State-Specific Solvent Exposure Analysis

Aβ40 monomers showed higher SASA values than Aβ42 monomers at D1, D7, E11, K16, F20, V24, S26, and at the CTR (Figure S3, right panel). Aβ42 monomers had higher SASA values than Aβ40 monomers at A2, in the F4-R5 region, V12, H14, N27, A30, I31, and M35. Amino acids with increased SASA values in Aβ40 relative to Aβ42 dimers were D1, H6, Y10, H13, and the G38-V40 region. Conversely, Aβ42 dimers showed higher SASA values than Aβ40 dimers in the A2-R5 region, at G9, K16-L17, E22, V24, K28, and in the A30-I31 region. The amino acids with higher SASA values in Aβ40 than in Aβ42 trimers were the H6-D7 region, Y10, H13, Q15, G38, and V40, whereas those with a higher solvent exposure in Aβ42 than in Aβ40 trimers were regions A2-F4 and F20-A21, and V36. Aβ40 tetramers had increased solvent exposure relative to Aβ42 tetramers at D1, H6, E11, F19, N27, and at the CTR, whereas Aβ42 tetramers displayed higher SASA values than Aβ40 tetramers at A2, R5, G9, H13, Q15, D23, K28, and in the M35-V36 region. The amino acids that were more exposed to the solvent in Aβ40 than Aβ42 pentamers were D1, H6, the H14-L17 region,

F19, V24, A30, M35, and the G37-V40 region, whereas those with higher SASA values in Aβ42 than in Aβ40 pentamers were the A2-E3 region, R5, the D7-G9 region, V12, E22, the N27-K28 and I31-L34 regions.

Figure S3 (right panels) shows the SASA per amino acid for monomers through pentamers of Aβ40 and Aβ42. We closely examined SASA changes in the NTR in particular and observed that Aβ40 and Aβ42 monomers showed the largest SASA values in the D1-Y10 region. The decrease in SASA values relative to Aβ40 monomers was noted in Aβ40 dimers at F4 and R5, in Aβ40 trimers in the D1-R5 region and at Y10, in Aβ40 tetramers at D1, in the F4-D7 region and at Y10, and in Aβ40 pentamers in the A2-S8 region and at Y10. Relative to Aβ42 monomers, SASA values decreased also in Aβ42 dimers at F4, H6, and Y10, in Aβ42 trimers in the D1-D7 region and at Y10, in Aβ42 tetramers in D1-A2 and F4-H6 regions and at Y10, and in Aβ42 pentamers in D1-A2 and F4-D7 regions and at Y10. These results indicate that the N-terminal region was involved in contact formation, which is consistent with increased frequency in quaternary contacts in Aβ40 and Aβ42 tetramers and pentamers (Figure 7, main manuscript).

A.2 Supplementary Tables

Table S1: Average values of NT-CM distance and hydrophobic SASA probability distributions. Distributions are displayed in Figures 3 and 4 in the main manuscript.

	NT-CM Distance [nm]		**Hydrophobic SASA [nm^2]**	
N	Aβ_{40}	Aβ_{42}	Aβ_{40}	Aβ_{42}
1	1.12	1.23	13.53	15.25
2	1.44	1.58	9.42	10.56
3	1.55	1.76	8.09	9.92
4	1.91	1.88	7.72	8.62
5	2.21	2.05	7.82	8.47

SEM values are less than 0.01.

Table S2: Intra-peptide salt bridge propensities for monomers. Each amino acid combination capable of forming salt bridges is represented. Salt-bridge propensities which are significantly different between Aβ40 and Aβ42 alloforms are highlighted in red.

Monomers	**Intra peptide [%]**	
AA Pair	Aβ_{40}	Aβ_{42}
R5 - D1	27 ± 6	37 ± 7
R5 - E3	24 ± 6	32 ± 6
R5 - D7	11 ± 4	6 ± 3
R5 - E11	18 ± 5	14 ± 5
R5 - E22	12 ± 4	12 ± 5
R5 - D23	4 ± 2	2 ± 2
Total R5	**95 ± 12**	**103 ± 12**
K16 - D1	4 ± 2	2 ± 2
K16 - E3	2 ± 2	5 ± 3
K16 - D7	3 ± 2	10 ± 4
K16 - E11	7 ± 3	8 ± 3
K16 - E22	6 ± 3	4 ± 3
K16 - D23	2 ± 2	5 ± 3
Total K16	**24 ± 6**	**35 ± 8**
K28 - D1	9 ± 3	6 ± 3
K28 - E3	3 ± 1	9 ± 4
K28 - D7	3 ± 2	2 ± 2
K28 - E11	5 ± 3	6 ± 3
K28 - E22	12 ± 4	15 ± 5
K28 - D23	14 ± 4	14 ± 4
Total K28	**46 ± 7**	**52 ± 9**

Table S3: Intra-peptide and inter-peptide salt bridge propensities for dimers. Each amino acid combination capable of forming salt bridges is represented. Salt-bridge propensities which are significantly different between Aβ40 and Aβ42 alloforms are highlighted in red.

Dimers	**Intra peptide [%]**		**Inter peptide [%]**		**Total [%]**	
AA Pair	Aβ_{40}	Aβ_{42}	Aβ_{40}	Aβ_{42}	Aβ_{40}	Aβ_{42}
R5 - D1	9 ± 3	28 ± 4	1 ± 1	0 ± 0	10 ± 3	28 ± 4
R5 - E3	31 ± 5	35 ± 4	0 ± 0	1 ± 1	31 ± 5	36 ± 4
R5 - D7	11 ± 4	16 ± 3	0 ± 0	0 ± 0	11 ± 4	16 ± 3
R5 - E11	6 ± 2	7 ± 3	6 ± 2	1 ± 1	11 ± 3	8 ± 3
R5 - E22	8 ± 3	4 ± 2	5 ± 3	6 ± 2	14 ± 4	10 ± 3
R5 - D23	3 ± 2	1 ± 1	3 ± 2	3 ± 2	5 ± 2	4 ± 2
Total R5	**67 ± 8**	**91 ± 8**	**15 ± 4**	**11 ± 3**	**82 ± 9**	**102 ± 8**
K16 - D1	7 ± 3	3 ± 1	2 ± 1	2 ± 1	9 ± 3	5 ± 2
K16 - E3	4 ± 2	4 ± 2	2 ± 1	3 ± 1	5 ± 2	7 ± 2
K16 - D7	4 ± 2	1 ± 1	1 ± 0	1 ± 1	4 ± 2	2 ± 2
K16 - E11	10 ± 2	4 ± 2	0 ± 0	3 ± 2	10 ± 2	8 ± 3
K16 - E22	4 ± 2	1 ± 1	2 ± 1	2 ± 1	6 ± 2	3 ± 1
K16 - D23	3 ± 1	2 ± 1	1 ± 1	3 ± 2	4 ± 2	5 ± 2
Total K16	**30 ± 5**	**14 ± 3**	**8 ± 2**	**15 ± 3**	**37 ± 5**	**28 ± 5**
K28 - D1	3 ± 1	5 ± 2	3 ± 2	1 ± 1	7 ± 2	6 ± 2
K28 - E3	4 ± 1	2 ± 1	5 ± 2	0 ± 0	8 ± 3	3 ± 1
K28 - D7	5 ± 2	2 ± 1	0 ± 0	2 ± 1	5 ± 2	3 ± 2
K28 - E11	2 ± 1	2 ± 1	1 ± 1	2 ± 1	4 ± 2	4 ± 2
K28 - E22	11 ± 3	7 ± 2	5 ± 2	5 ± 2	16 ± 3	12 ± 3
K28 - D23	17 ± 4	10 ± 3	3 ± 1	2 ± 1	20 ± 4	12 ± 3
Total K28	**41 ± 5**	**27 ± 5**	**18 ± 4**	**12 ± 3**	**59 ± 7**	**39 ± 6**

Table S4: Intra-peptide and inter-peptide salt bridge propensities for trimers. Each amino acid combination capable of forming salt bridges is represented. Salt-bridge propensities which are significantly different between Aβ40 and Aβ42 alloforms are highlighted in red.

Trimers	**Intra peptide [%]**		**Inter peptide [%]**		**Total [%]**	
AA Pair	Aβ_{40}	Aβ_{42}	Aβ_{40}	Aβ_{42}	Aβ_{40}	Aβ_{42}
R5 - D1	16 ± 4	34 ± 6	3 ± 1	0 ± 0	19 ± 5	34 ± 6
R5 - E3	27 ± 5	28 ± 5	2 ± 1	2 ± 1	28 ± 5	28 ± 5
R5 - D7	12 ± 3	14 ± 4	1 ± 1	1 ± 1	12 ± 3	16 ± 4
R5 - E11	7 ± 3	17 ± 4	3 ± 1	1 ± 1	11 ± 3	19 ± 4
R5 - E22	9 ± 3	6 ± 3	3 ± 1	2 ± 1	12 ± 3	7 ± 4
R5 - D23	6 ± 3	5 ± 3	4 ± 1	2 ± 1	10 ± 3	7 ± 3
Total R5	**78 ± 9**	**104 ± 11**	**15 ± 3**	**8 ± 2**	**93 ± 9**	**112 ± 11**
K16 - D1	6 ± 2	2 ± 1	2 ± 1	4 ± 1	8 ± 3	5 ± 2
K16 - E3	2 ± 1	3 ± 2	3 ± 1	3 ± 1	5 ± 2	5 ± 2
K16 - D7	1 ± 1	7 ± 2	0 ± 0	0 ± 0	2 ± 1	7 ± 2
K16 - E11	3 ± 2	5 ± 2	2 ± 1	2 ± 1	5 ± 2	7 ± 2
K16 - E22	4 ± 2	2 ± 2	2 ± 1	2 ± 1	6 ± 2	4 ± 2
K16 - D23	4 ± 2	3 ± 2	0 ± 0	3 ± 1	4 ± 2	6 ± 2
Total K16	**20 ± 4**	**20 ± 4**	**9 ± 2**	**14 ± 3**	**28 ± 5**	**34 ± 5**
K28 - D1	5 ± 3	1 ± 0	2 ± 1	3 ± 1	6 ± 3	3 ± 1
K28 - E3	2 ± 1	2 ± 2	2 ± 1	2 ± 1	3 ± 2	4 ± 2
K28 - D7	2 ± 2	2 ± 2	1 ± 1	1 ± 0	3 ± 2	3 ± 2
K28 - E11	4 ± 2	3 ± 2	1 ± 1	2 ± 1	5 ± 2	4 ± 2
K28 - E22	7 ± 2	12 ± 4	3 ± 1	2 ± 1	10 ± 3	14 ± 4
K28 - D23	15 ± 4	11 ± 3	2 ± 1	2 ± 1	17 ± 4	13 ± 3
Total K28	**33 ± 6**	**30 ± 6**	**11 ± 3**	**10 ± 2**	**44 ± 6**	**40 ± 6**

Table S5: Intra-peptide and inter-peptide salt bridge propensities for tetramers. Each amino acid combination capable of forming salt bridges is represented. Salt-bridge propensities which are significantly different between Aβ40 and Aβ42 alloforms are highlighted in red.

Tetramers	**Intra peptide [%]**		**Inter peptide [%]**		**Total [%]**	
AA Pair	Aβ_{40}	Aβ_{42}	Aβ_{40}	Aβ_{42}	Aβ_{40}	Aβ_{42}
R5 - D1	22 ± 4	52 ± 7	3 ± 1	1 ± 1	25 ± 4	53 ± 7
R5 - E3	22 ± 5	22 ± 5	1 ± 1	0 ± 0	24 ± 5	22 ± 5
R5 - D7	13 ± 4	11 ± 4	1 ± 1	2 ± 1	13 ± 4	13 ± 5
R5 - E11	12 ± 4	1 ± 1	2 ± 1	1 ± 1	15 ± 4	2 ± 1
R5 - E22	8 ± 3	0 ± 0	1 ± 1	3 ± 2	9 ± 4	3 ± 2
R5 - D23	6 ± 2	3 ± 2	0 ± 0	0 ± 0	6 ± 2	3 ± 2
Total R5	**83 ± 9**	**90 ± 10**	**8 ± 2**	**8 ± 2**	**91 ± 10**	**97 ± 10**
K16 - D1	1 ± 0	0 ± 0	2 ± 1	2 ± 1	3 ± 1	2 ± 1
K16 - E3	0 ± 0	1 ± 1	2 ± 1	2 ± 1	2 ± 1	3 ± 2
K16 - D7	6 ± 2	3 ± 2	1 ± 1	1 ± 1	7 ± 3	4 ± 2
K16 - E11	4 ± 2	4 ± 2	1 ± 1	1 ± 1	5 ± 2	5 ± 2
K16 - E22	1 ± 1	5 ± 3	2 ± 1	2 ± 1	2 ± 1	7 ± 3
K16 - D23	2 ± 1	1 ± 1	1 ± 1	2 ± 1	3 ± 2	2 ± 1
Total K16	**13 ± 3**	**14 ± 4**	**9 ± 2**	**9 ± 2**	**21 ± 4**	**22 ± 5**
K28 - D1	3 ± 1	2 ± 2	2 ± 1	1 ± 1	5 ± 2	3 ± 2
K28 - E3	3 ± 2	0 ± 0	0 ± 0	1 ± 0	4 ± 2	1 ± 0
K28 - D7	1 ± 1	2 ± 1	1 ± 1	1 ± 1	2 ± 1	3 ± 2
K28 - E11	1 ± 0	0 ± 0	0 ± 0	1 ± 1	1 ± 1	1 ± 1
K28 - E22	11 ± 4	14 ± 3	4 ± 1	1 ± 0	15 ± 4	15 ± 3
K28 - D23	22 ± 5	5 ± 2	3 ± 1	2 ± 1	25 ± 5	7 ± 2
Total K28	**41 ± 7**	**24 ± 4**	**10 ± 2**	**6 ± 1**	**51 ± 7**	**30 ± 4**

Table S6: Intra-peptide and inter-peptide salt bridge propensities for pentamers. Each amino acid combination capable of forming salt bridges is represented. Salt-bridge propensities which are significantly different between Aβ40 and Aβ42 alloforms are highlighted in red.

Pentamers	**Intra peptide [%]**		**Inter peptide [%]**		**Total [%]**	
AA Pair	Aβ_{40}	Aβ_{42}	Aβ_{40}	Aβ_{42}	Aβ_{40}	Aβ_{42}
R5 - D1	8 ± 3	20 ± 2	1 ± 1	2 ± 1	9 ± 3	23 ± 2
R5 - E3	20 ± 5	22 ± 7	1 ± 1	2 ± 1	21 ± 5	23 ± 7
R5 - D7	1 ± 1	14 ± 6	3 ± 1	1 ± 1	4 ± 1	15 ± 6
R5 - E11	2 ± 2	10 ± 5	3 ± 1	1 ± 1	5 ± 2	11 ± 5
R5 - E22	0 ± 0	0 ± 0	3 ± 2	2 ± 1	3 ± 2	2 ± 1
R5 - D23	4 ± 3	0 ± 0	2 ± 1	1 ± 1	6 ± 4	1 ± 1
Total R5	**35 ± 7**	**65 ± 11**	**13 ± 3**	**9 ± 3**	**48 ± 8**	**74 ± 11**
K16 - D1	7 ± 3	4 ± 4	1 ± 1	2 ± 1	8 ± 3	7 ± 4
K16 - E3	3 ± 3	1 ± 1	2 ± 1	2 ± 1	5 ± 3	2 ± 1
K16 - D7	0 ± 0	3 ± 2	1 ± 1	1 ± 1	1 ± 1	4 ± 2
K16 - E11	9 ± 5	10 ± 3	1 ± 1	0 ± 0	10 ± 5	10 ± 3
K16 - E22	0 ± 0	1 ± 1	1 ± 1	3 ± 1	1 ± 1	3 ± 1
K16 - D23	0 ± 0	0 ± 0	1 ± 1	1 ± 1	1 ± 1	1 ± 1
Total K16	**19 ± 6**	**18 ± 5**	**8 ± 2**	**9 ± 2**	**27 ± 7**	**27 ± 6**
K28 - D1	0 ± 0	6 ± 3	3 ± 1	1 ± 0	3 ± 1	7 ± 3
K28 - E3	3 ± 3	3 ± 2	1 ± 1	1 ± 1	4 ± 3	4 ± 2
K28 - D7	4 ± 3	1 ± 0	2 ± 1	1 ± 1	6 ± 4	1 ± 1
K28 - E11	3 ± 3	0 ± 0	1 ± 0	1 ± 1	3 ± 3	1 ± 1
K28 - E22	9 ± 4	9 ± 7	3 ± 1	1 ± 1	12 ± 4	10 ± 7
K28 - D23	17 ± 2	9 ± 5	4 ± 2	0 ± 0	21 ± 3	9 ± 5
Total K28	**36 ± 7**	**27 ± 9**	**12 ± 2**	**6 ± 1**	**48 ± 7**	**33 ± 9**

A.3 Supplementary Figures

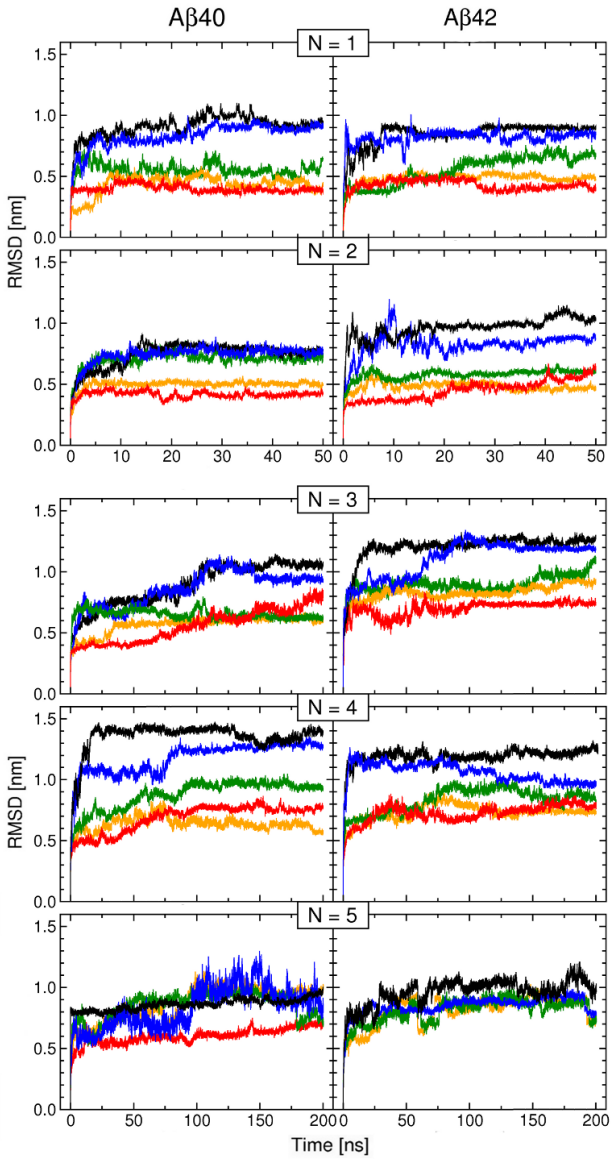

Figure S1: Temporal evolution of RMSD values. RMSD values as a function of simulation time for trajectories of Aβ40 and Aβ42 monomers and oligomers obtained by MD. Trajectories with the highest and lowest RMSD values were chosen along with three representative trajectories. RMSD values for all four Aβ42 pentamers are shown.

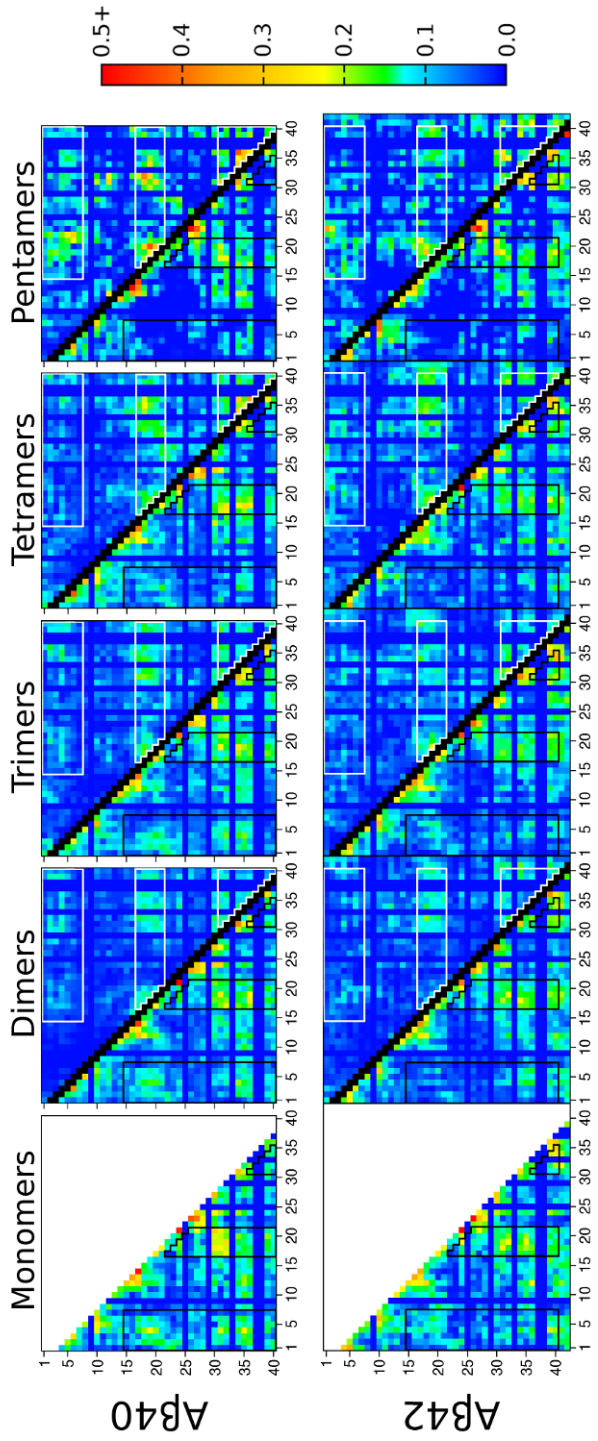

Figure S2: Cβ–Cβ contact maps. Intramolecular contacts are below the black diagonal line, intermolecular contacts are above. The color scale on the right shows the contact propensity as defined in *Methods*. SEM values for all contact propensities were < 0.05 (data not shown).

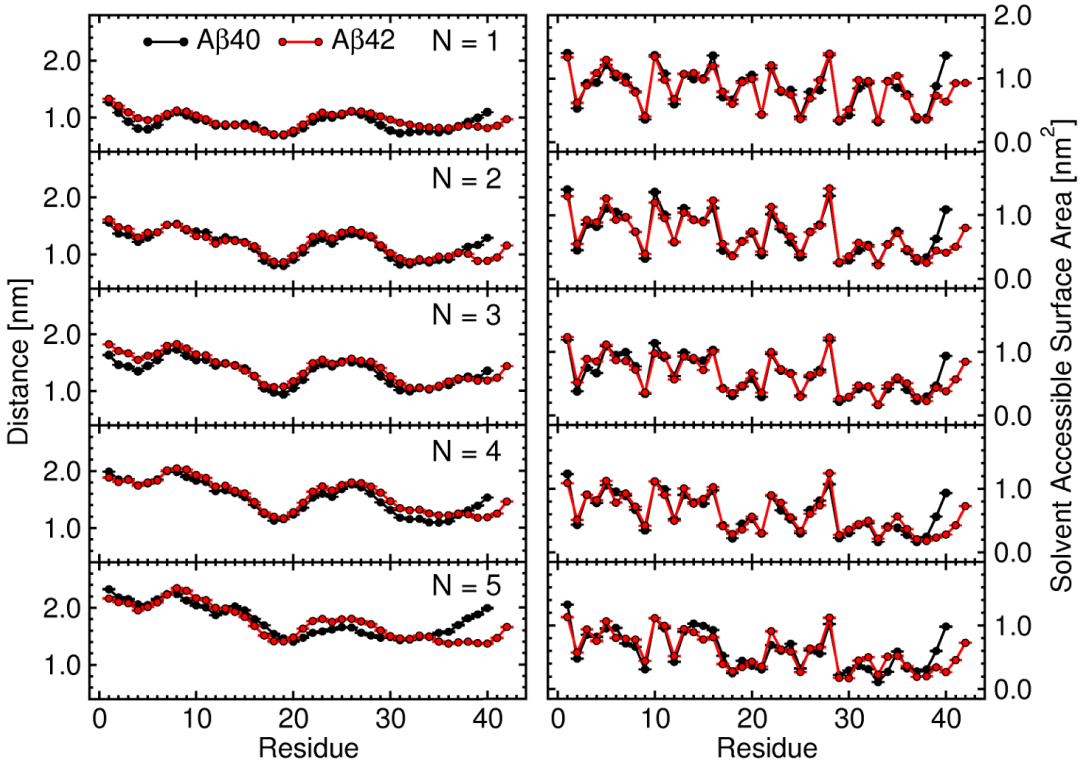

Figure S3: Distance from the CM and SASA per amino acid: alloform comparison. The error bars are SEM values.

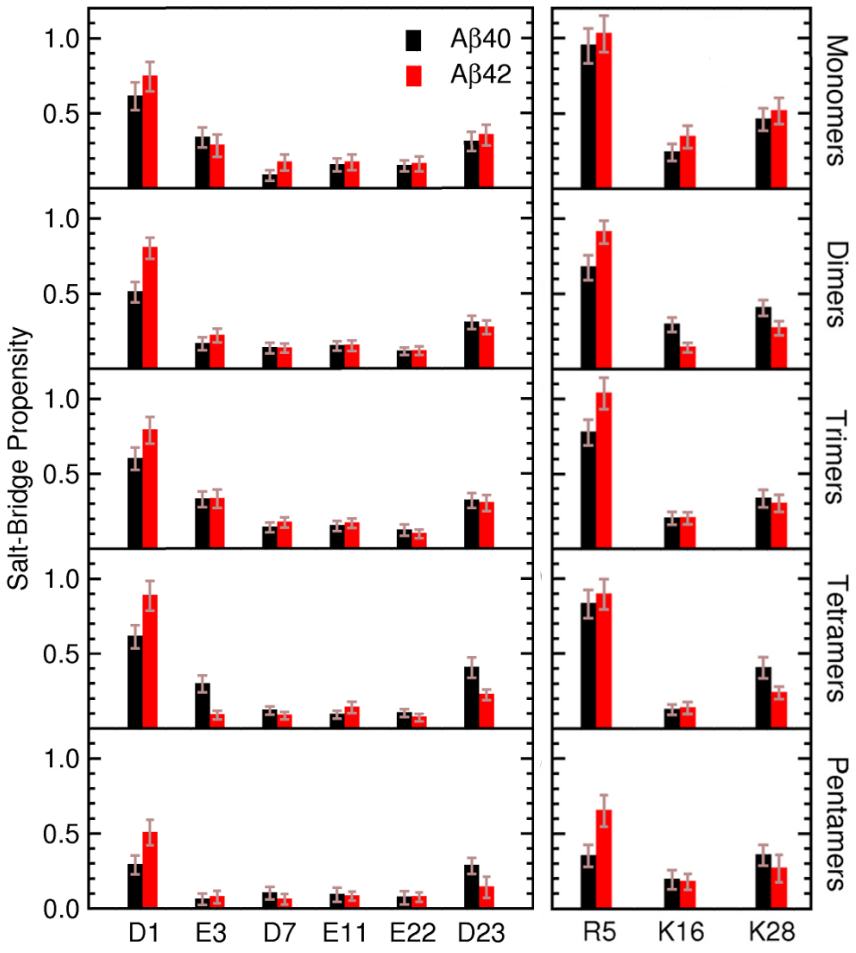

Figure S4: Histograms of salt-bridge propensities: alloform effect. Sums of intra and inter-peptide propensities, for the three positive (right panel) and six negative (left panel) amino acids are shown. The error bars are SEM values.

Appendix B: Supporting Information for: *Folding and Oligomer Formation of α-Synuclein: A Discrete Molecular Dynamics Study*

B.1 Supplemental Tables

Table S1: Number of monomer and oligomer conformations. Throughout the 80M time units of simulation time and in the final 20M time units of simulation time. Specific time frames used for analysis are discussed in the *Methods* section.

N	0-80M	60-80M
1	107,351	19,766
2	35,398	7,079
3	18,473	3,938
4	9,940	2,499
5	6,298	1,624
8	1,821	862

B.2 Supplemental Figures

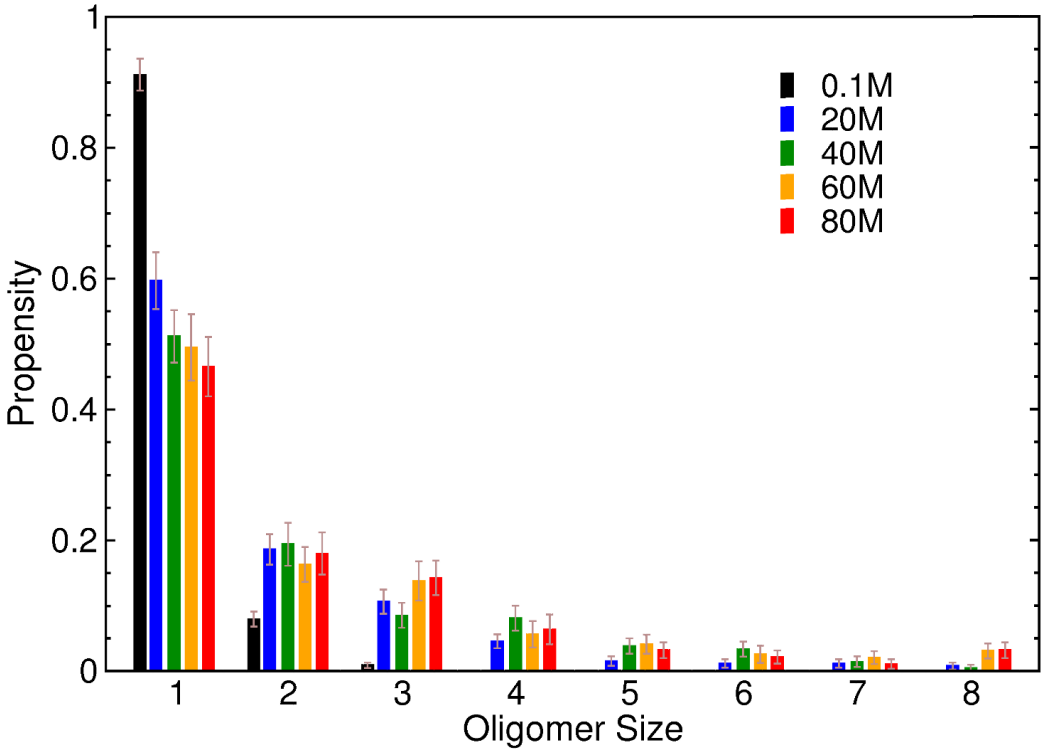

Figure S1: Time evolution of the oligomer size distribution. The relative propensity of monomers and oligomers for simulations with $E_{CH} = 0.20$ was measured after the first 1M time units and at 20M, 40M, 60M, and 80M time units.

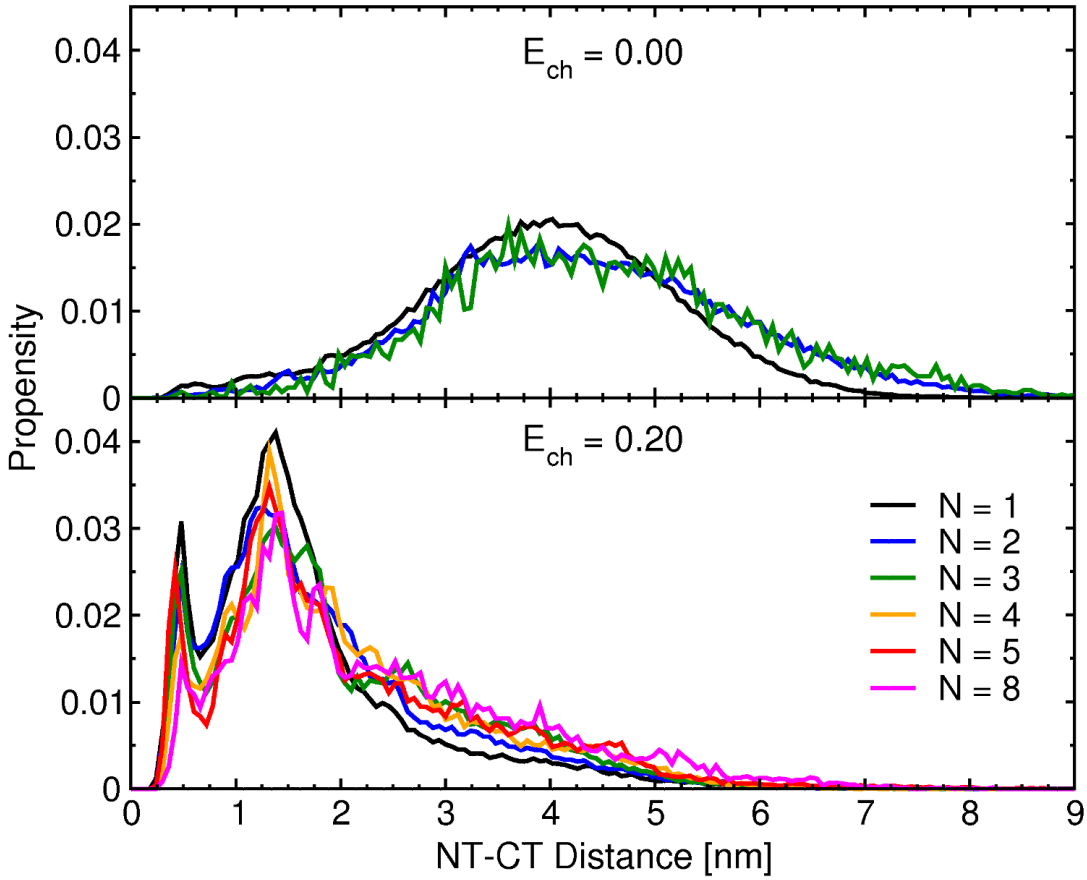

Figure S2: Effect of assembly state on distributions of NT-CM, CT-CM, and NT-CT distances and hydrophobic CG-SASA. Mann-Whitney U tests comparing all assembly size distributions resulted in $p < 10^{-3}$ for all.

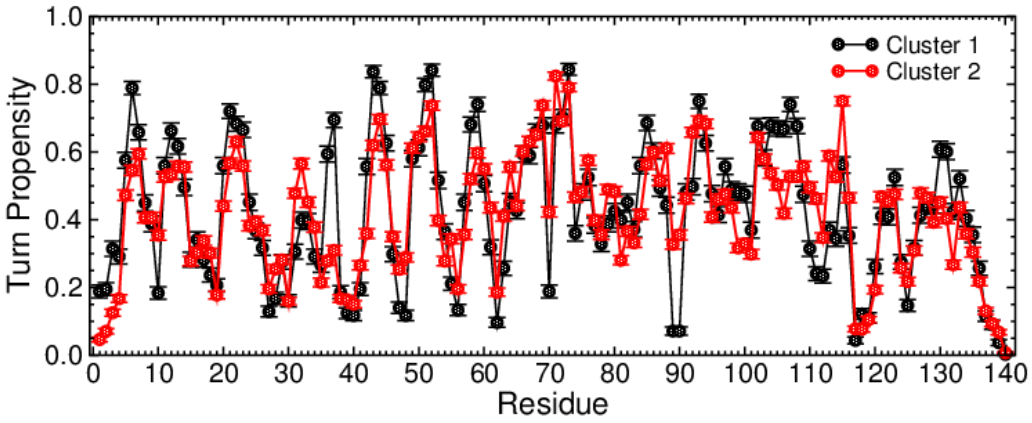

Figure S3: Average per-residue turn propensities for monomer conformations from clusters 1 and 2. Error bars correspond to SEM values

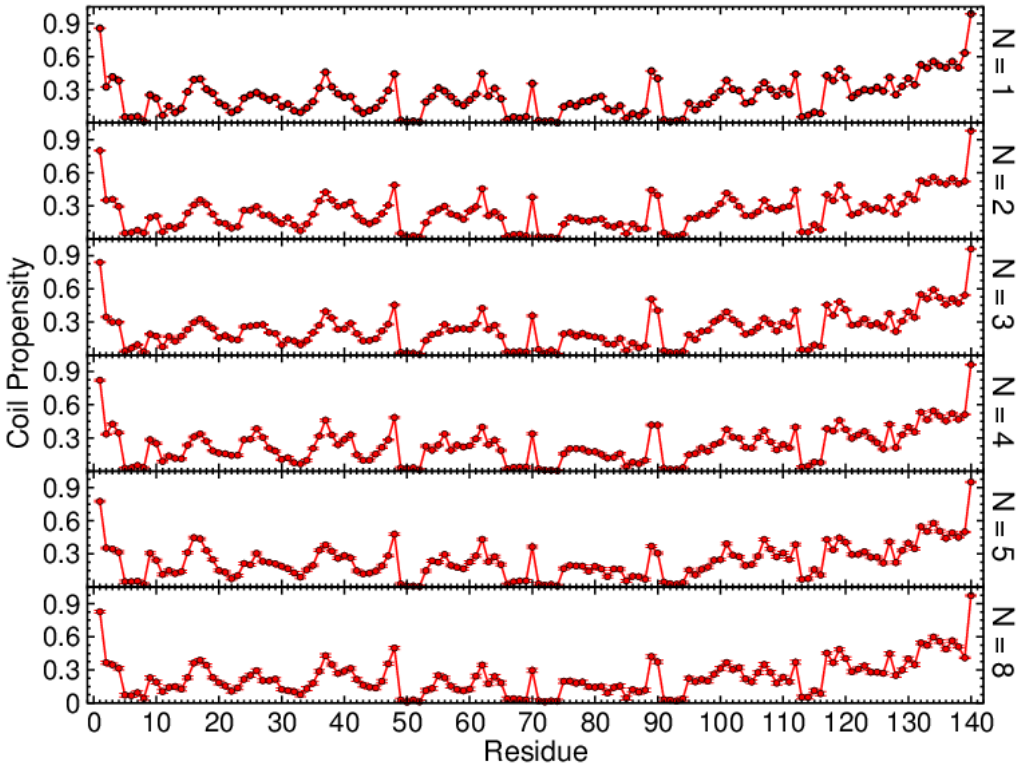

Figure S4: Average per-residue statistical coil content of αS conformations. Error bars correspond to SEM values.

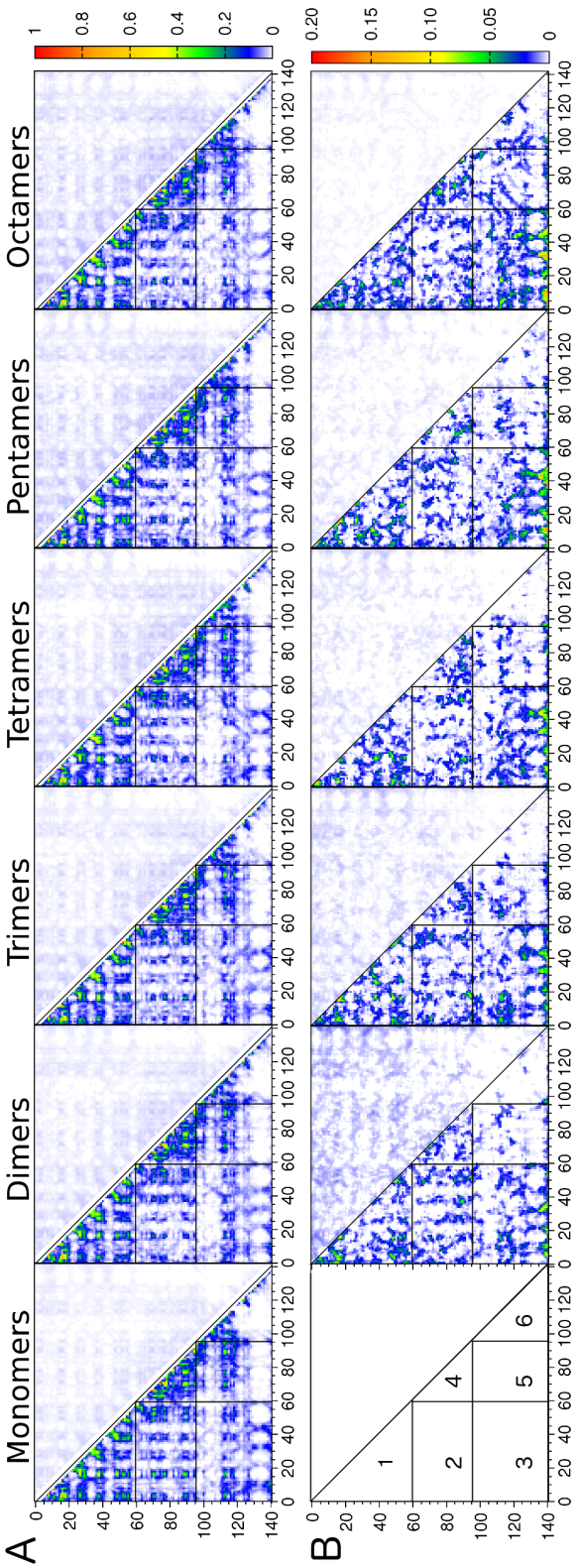

Figure S5: C_α-C_α contact maps. Intramolecular (A) and intermolecular (B) contacts are below the diagonal with corresponding SEM values above the diagonal. The color scale on the right shows the contact propensity as defined in the *Methods* section. Contact maps are divided into 6 sectors (bottom left).

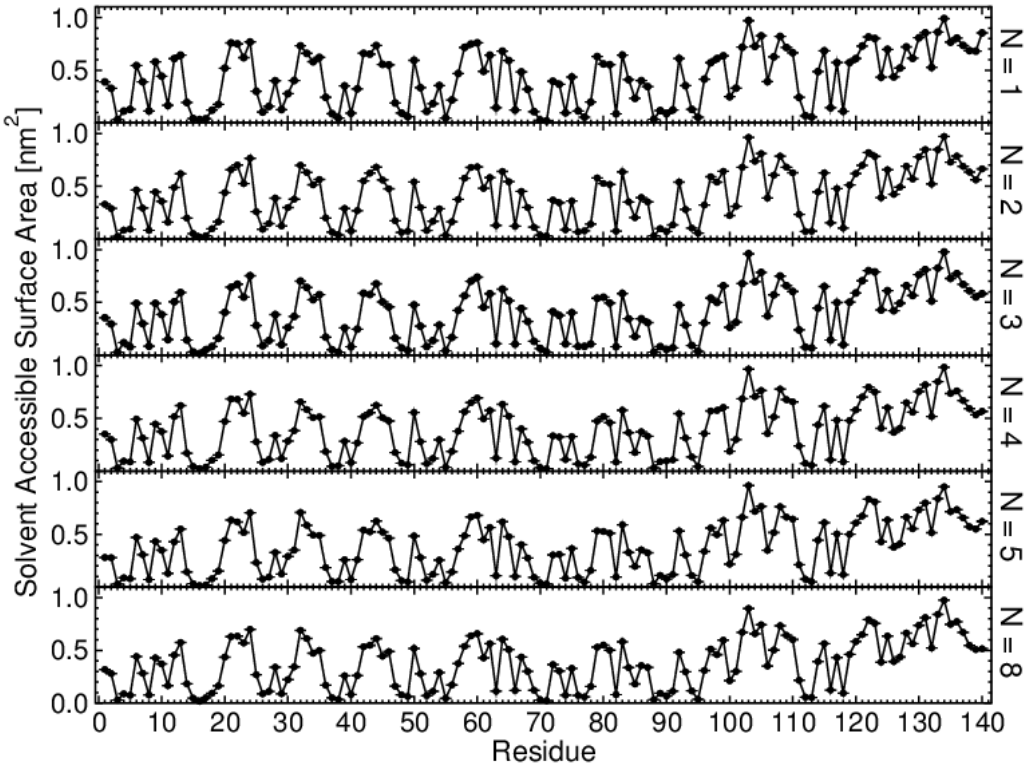

Figure S6: Average per-residue SASA of αS conformations. Error bars correspond to SEM values.